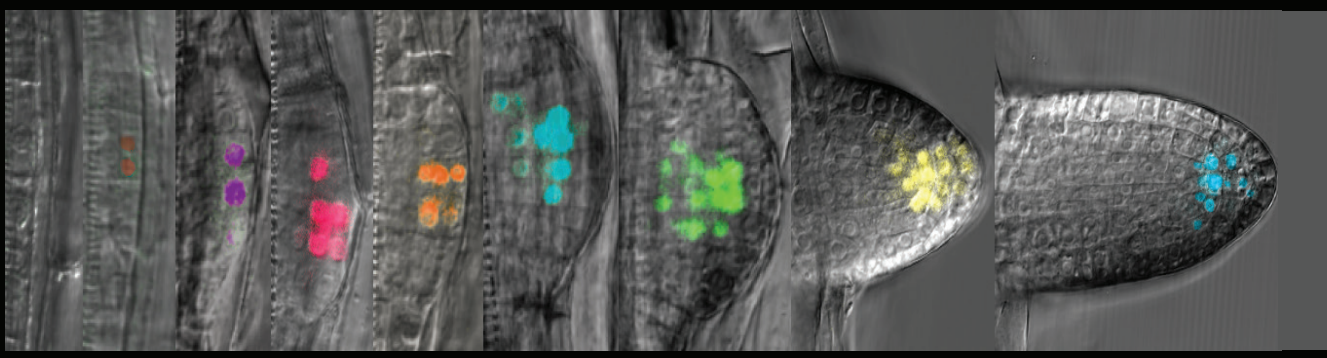
Plant signaling peptides, novel insights into their processing and role in root development

Sarieh Ghorbani



2014

Cover: *GLV10* expression pattern during lateral root primordia development

To my wonderful parents, Samarrokh and Bohloul

تقدیم به مهربانی ها و همراهی های پدر و مادر عزیزم، ثمرخ و بهلول

یک چند به کودکی به استاد شدیم　　یک چند به استادی خود شاد شدیم

پایان سخن شنو که ما را چه رسید　　از خاک بر آمدیم و بر باد شدیم

حکیم عمر خیام

Myself, when young did eagerly frequent
Doctor and Saint, and heard great Argument
About it and about: but evermore
Came out by the same Door where in I went.

Omar Khayyam (Persian poet, mathematician, astronomer, 1048 –1131)
(translated by Edward FitzGerald)

Ghent University
Faculty of Sciences
Department of Plant Biotechnology and Bioinformatics

Plant signaling peptides, novel insights into their processing and role in root development

Sarieh Ghorbani

Promoters: **Prof. Dr. Tom Beeckman** and **Dr. Pierre Hilson**

April 2014

VIB – Plant Systems Biology

Thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Science: Biochemistry and Biotechnology

Promoters:

Prof. Dr. Tom Beeckman

Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University
tobee@psb.vib-ugent.be

Dr. Pierre Hilson

Institut Jean-Pierre Bourgin, UMR1318 INRA-AgroParisTech, Versailles Cedex, France
Pierre.Hilson@versailles.inra.fr

Examination commission:

Prof. Dr. Ann Depicker (Chair)

Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University
anpic@psb.vib-ugent.be

Prof. Dr. Rüdiger Simon (reading committee)

Institut für Genetik, Heinrich-Heine Universität, Düsseldorf, Germany
ruediger.simon@hhu.de

Prof. Dr. Frank Van Breusegem (reading committee)

Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University
frbre@psb.vib-ugent.be

Prof. Dr. Annemieke Madder

Department of Organic Chemistry, Laboratory for Organic and Biomimetic Chemistry, Ghent University
annemieke.madder@ugent.be

Prof. Dr. Els Vandamme (reading committee)

Department of Molecular Biotechnology, Faculty of Bioscience Engineering, Ghent University
ElsJM.VanDamme@ugent.be

Dr. Ana Fernandez

Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University
anfer@psb.vib-ugent.be

Dr. Laurens Pauwels (reading committee)

Department of Plant Systems Biology, VIB and Department of Plant Biotechnology and Bioinformatics, Ghent University
lapau@psb.ugent.be

TABLE OF CONTENTS

Frequently used abbreviations

ACR4:	*ARABIDOPSIS CRINKLY4*
APC/C:	*ANAPHASE-PROMOTING COMPLEX/CYCLOSOME*
ARF7:	*AUXIN RESPONSIVE FACTOR7*
BiFC:	Bimolecular fluorescence complementation
CC:	Columella cells
CCS52A2:	*CELL CYCLE SWITCH 52A2*
CLE:	*CLAVATA3/ESR-RELATED*
CHA*L*:	*CHALLAH*
CLV:	*CLAVATA*
COFRADIC:	Combined fractional diagonal chromatography
CPD:	*CARBOXYPEPTIDASE D*
CPE:	*CARBOXYPEPTIDASE E*
CRN:	*CORYNE*
CEP:	*C-TERMINALLY ENCODED PEPTIDE*
CT2:	*COMPACT PLANT2*
DAG:	Day after germination
ENOD40:	*EARLY NODULIN40*
EPF:	*EPIDERMAL PATTERNING FACTOR*
ER:	Endoplasmic reticulum
ER:	*ERECTA*
ERF115:	*ETHYLENE RESPONSIVE FACTOR 115*
ERL:	*ERECTA-LIKE*
ESR:	*ENDOSPERM SURROUNDING REGION*
FEA2:	*FASCIATED EAR2*
FRET:	Fluorescence resonance energy transfer
GFP:	Green fluorescent protein
GLV/RGF/CLEL:	*GOLVEN/ROOT GROWTH FACTOR/CLE-LIKE*
GUS:	β-glucuronidase
HPAT	*HYDROXYPROLINE O-ARABINOSYLTRANSFERASE*
IAA:	Indole-3-acetic acid
IDA:	*INFLORESCENCE DEFICIENT IN ABSCISSION*
IDL:	*IDA-LIKE*
LAX3:	*auxin influx carrier LIKE AUX1-3*
LR:	Lateral root
LRD:	Lateral root density
LRI:	Lateral root initiation
LRP:	Lateral root primordium
LRR-RLK:	Leucine-rich repeat-receptor-like kinase
MC9:	*METACASPASE 9*
NLS:	Nuclear localization signal
OE:	Overexpression
P4H:	*PROLYL 4-HYDROXYLASE*
PI:	Propidium iodide
PIN2:	*PIN-FORMED2*
PLT:	*PLETHORA*
PSK:	*PHYTOSULFOKINE*
PSY:	*PLANT PEPTIDE CONTAINING SULFATED TYROSINE*

PXY:	*PHLOEM INTERCALATED WITH XYLEM*
QC:	Quiescent center
qRT-PCR:	Quantitative reverse transcriptase-polymerase chain reaction
RAM:	Root apical meristem
RALF:	*RAPID ALKALINIZATION FACTOR*
RD21:	*RESPONSIVE-TO-DESICCATION-21*
RFP:	Red fluorescent protein
SAM:	Shoot apical meristem
SBT:	*SUBTILISIN-LIKE SERINE PROTEASE*
SRP:	*SERPIN*
SSP:	Small signaling peptide
SPP:	*SIGNAL PEPTIDE PEPTIDASE*
TDR:	*TDIF RECEPTOR*
TF:	Transcription factor
TMM:	*TOO MANY MOUTHS*
TPPII:	*TRIPEPTIDYL PEPTIDASE II*
TPST:	*TYROSYL PROTEIN SULFOTRANSFERASE1*
TDIF:	*TRACHEARY ELEMENT DIFFERENTIATION INHIBITORY FACTOR*
WT:	Wild-type
WUS:	*WUSCHEL*
WOX:	*WUSCHEL-related homeobox*

Summary

Growth and development in multicellular organisms are highly dependent on cell-to-cell communication. Although the importance of signaling peptides in cell-to-cell communication networks is well acknowledged to date, identification and functional characterization of these molecules is, however, not very advanced yet. The majority of secreted peptides are encoded by small genes and therefore tend to be missed in genome annotations. Based on a few structural features that characterize small signaling peptides known so far and expression pattern analysis at specific developmental stages and in specific cell types, we identified novel genes coding for previously uncharacterized oligopeptides, possibly involved in cell-to-cell communication during lateral root (LR) formation in *Arabidopsis thaliana* and other plant species.

One of the new peptide families identified by this *in silico* screening is the *GOLVEN* (*GLV*) gene family which encodes small secreted peptides involved in important plant developmental programs including root and LR development. Although GLV peptides have been partially characterized already, nothing is known about the factors required for the production of the mature bioactive GLV peptides. Through a genetic suppressor screen in *Arabidopsis thaliana*, we identified two related subtilase genes, *AtSBT6.1* and *AtSBT6.2*, necessary for GLV1 activity. Root and hypocotyl *GLV1* overexpression phenotypes were suppressed by mutations in either subtilase gene. Synthetic GLV-derived peptides were cleaved *in vitro* by the affinity-purified SBT6.1 catalytic enzyme, confirming that the GLV1 precursor is a direct subtilase substrate. Furthermore, the protease inhibitor Serpin1 bound to SBT6.1 and inhibited the cleavage of GLV1 precursors by the protease. In agreement with the role of the SBT6 subtilases in GLV precursor processing, both null mutants for *sbt6.1* and *sbt6.2* and the *Serpin1* overexpression plants had shorter hypocotyls.

Our data fits with a model in which the GLV1 signaling pathway participates in the regulation of hypocotyl cell elongation, which is controlled by SBT6 subtilases, and modulated locally by the Serpin1 protease inhibitor.

Additional studies in lateral root development indicate that different *GLV* genes contribute to primordium organogenesis and that GLV10 regulates lateral root primordium development. Furthermore, our data indicates that GLV10 might be involved in root apical meristem maintenance, together with the previously reported GLV5, GLV7, and GLV11 peptides.

Scope

In multicellular organisms, growth and development need to be precisely coordinated, strongly relying on positional information, which is achieved through exchanges of molecular messages between cells and tissues by means of cell-to-cell communication mechanisms. Especially in plants, accurate and well-controlled cell-to-cell communication networks are essential, because of the complete absence of cell mobility and the presence of rigid cell walls. For many years, phytohormones had been thought to be the only communication messengers. Identification of systemin as the first plant signaling peptide revealed that counterparts of the mammalian peptide hormones were present in plants as well. Although over the past years several plant signaling peptides have been discovered and implicated in different growth and developmental aspects, the molecular basis of the plant peptide signaling mechanisms remains largely unknown.

The recently identified signaling peptide family, designated GOLVEN (GLV), has been shown to be involved in a broad range of developmental processes such as root apical meristem maintenance, gravitropic response, lateral root and root hair development in *Arabidopsis thaliana*. However, many questions related to its production, activity and signal transduction pathways are still unanswered.

Signaling peptides, including the GLV family, are generally encoded as large precursors that need to undergo several posttranslational modifications and processing events to yield mature peptides. In order to have a better insight into the maturation and activation of the GLV signaling peptides, we aimed at finding the potential enzyme(s) responsible for processing of GLV precursor(s) as a first objective of this PhD project. GLV precursors have similar structure to subtilase targets, suggesting that they are possibly recognized and processed by these proteases. This prompted us to screen an available *Arabidopsis* subtilase

mutant collection to explore the possibility that any of these family members is functioning in GLV signaling.

In the second part, we focused on the significance of signaling peptides for root development. The *Arabidopsis* genome contains thousands of small open reading frames potentially coding for small signaling peptides, of which some might possibly be involved in root development. However, genes encoding small peptides tend to be absent in genome annotations, because it is difficult to distinguish them from short, random open reading frames. Additionally, detection of the mature small signaling peptide from crude extracts of plant tissues is not always an easy task given their very low physiological concentration. Therefore, it is likely that only a few of the secreted peptides are known to date. To this end, we first performed an *in silico* analysis to screen for new signaling peptides with an increased probability to be involved in lateral root development of *Arabidopsis*.

GLV peptides have been suggested to be involved in numerous developmental features, including root and LR development, we focused also on the characterization of the GLV signaling peptide family during primary root and LR development. One of the *GLV* genes expressed early during LR formation is *GLV10.* Finally, we investigated the role of GLV10 in root development for which evidence for a potential root function had been suggested by its expression pattern.

Chapter 1

Introduction to root development and signaling peptides

Introduction

I. Root development

The root system is fundamentally important for plants because of its diverse functions, such as mineral acquisition, water uptake, anchoring, and also production and storage of various metabolites. Due to the sessile nature of plants, the developmental plasticity of the root system is a major characteristic enabling plants to adapt to different environments (Fitter, 2002; Guyomarc'h, 2010). The *Arabidopsis* root is a valuable system for developmental studies thanks to its simple morphological and structural organization (Dolan et al., 1993; Van Norman et al., 2013a).

I.1. Primary root structure

In the majority of plants, the primary root (derived from the embryonic root or radical) is the first structure that emerges from the germinating seed (Clowes, 1961). In *Arabidopsis* roots the root cap (RC) is located at the distal tip of the root, composed of columella cells (CCs) and lateral root cap cells. Just above the CCs, a quiescent center (QC) is located from where the main body of the root starts (Figure 6A). The QC, together with the surrounding stem cells, are called the stem cell niche (Dolan et al., 1993; Barlow, 2002; Gallagher, 2013).

The main body of the root consists of distinct tissue layers that contribute to different functional aspects. In *Arabidopsis*, the tissue layers from the outside to the inside are epidermis, cortex, endodermis, pericycle and finally vascular bundle. The vascular bundle contains vascular elements and consists of xylem and phloem cells (Dolan et al., 1993). The root apex is divided into different zones based on the shape and behavior of the cells. The meristematic zone is defined by active cell divisions and can itself be divided into two zones: (i) the apical meristem, where cells divide most actively and (ii) the basal meristem; which refers to the region between the root apical meristem (RAM) and the differentiation zone, it is a transition zone in the root tip where cell division stops and cells start to elongate (Dolan et al., 1993; Baluska et al., 2005; De Smet et al., 2007; Bennett and Scheres, 2010) (Figure 6B). Afterwards, when cells exit the cell cycle and leave the meristematic zone, they enter into the elongation zone to undergo elongation (Baluska et al., 2010) and finally, reach the differentiation zone, which is associated with the final differentiation into different cell types,

as reflected by the deposition of a secondary cell wall in xylem cells and the appearance of the root hairs (Bennett and Scheres, 2010; Perilli et al., 2012).

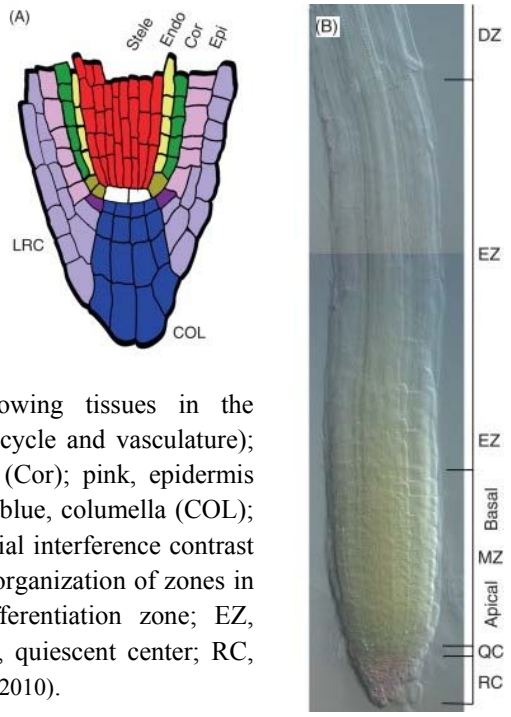


Figure 6: (A) Schematic diagram showing tissues in the *Arabidopsis* root meristem; red, stele (pericycle and vasculature); yellow, endodermis (endo); green, cortex (Cor); pink, epidermis (Epi); light purple, lateral root cap (LRC); blue, columella (COL); white, quiescent center (QC). (B) Differential interference contrast light micrograph showing the longitudinal organization of zones in the *Arabidopsis* root meristem; DZ, differentiation zone; EZ, elongation zone; MZ, meristem zone; QC, quiescent center; RC, root cap, (adapted from Bennett T. and Scheres 2010).

I.2. Lateral root development

Post-embryonically, plant growth depends strongly on the establishment of a root branching system to tap more remote resources of nutrients and water (Lavenus et al., 2013). Environmental conditions can induce or repress the development of LRs and, therefore, provide a mechanism by which plants can cope with changing soil conditions (Malamy, 2005).

LRs generally arise from pericycle cells, either adjacent to xylem poles in dicotyledonous plants, such as *Arabidopsis*, or phloem poles in some monocotyledonous plants (Dolan et al., 1993; Charlton, 1996; Osmont et al., 2007). Detailed studies in *Arabidopsis* have revealed the existence of several distinct developmental stages such as oscillation of gene expression, prebranch site formation and priming of pericycle cells, founder cell specification, lateral root initiation (LRI), lateral root primordia (LRP) development and LRP emergence (Malamy and Benfey, 1997; De Rybel et al., 2010; Dubrovsky, 2012; Van Norman et al., 2013b).

Even though LRs become evident first in the mature part of the main root, the primary phases of LR formation are thought to start far below the emergence site in the basal meristem or transition zone. Pericycle cells are arranged into longitudinal cell files that span the entire length of the root; nonetheless, only a few cells achieve the founder cell identity (Van Norman et al., 2013b). An oscillatory gene transcription system is considered the first known event in LR formation (Figure 7). Oscillation occurs in the basal meristem and elongation zone and is required for the establishment of the spatiotemporal distribution of LRs along the primary root (De Smet et al., 2007; Moreno-Risueno et al., 2010). The oscillation event is traceable using DR5 promoter activity reporting transcriptional responses to auxin. Although several hormones play a role in LR development, auxin signaling has a significant impact at nearly all the stages (Jansen et al., 2013). Only a few pericycle cells exposed to the oscillation will become founder cells. (Van Norman et al., 2013b).

The xylem pole pericycle cells that will give rise to LRP become specified as founder cells within a "prebranch site", a region of the root above the oscillation zone that is marked with static DR5-luciferase expression (Van Norman et al., 2013b). Although no cellular resolution can be obtained with the luciferase construct, an auxin response maximum could be detected in protoxylem cells in the oscillation zone by means of a DR5-β-glucuronidase (GUS) construct. Therefore, it has been hypothesized that the auxin response maximum shifts from the protoxylem cells to the pericycle founder cells in the prebranch sites. Interestingly, the expression of the transcription factor GATA23 seems to coincide with this protoxylem-to-pericycle shift of the auxin response maximum and occurs in stretches of the basal meristem-leaving pericycle cells. The frequency of this induction correlates with the periodicity of the DR5 oscillation (De Rybel et al., 2010). The take-off of the GATA23 expression in the pericycle precedes the DR5 expression and, thus, the founder cell specification and is interpreted as a preparative phase towards LRI, which was designated as "priming" (De Rybel et al., 2010). It is thought that during all these developmental phases, after leaving the RAM, the phloem pole pericycle cells remain in the G1 phase, whereas the xylem pole pericycle cells shift to the G2 phase, making them competent to divide (Beeckman et al., 2001).

Founder cell specification is almost immediately followed by asymmetric cell division and the start of a new LR. The LRI phase is recognizable by migration of the nuclei of two founder cells towards the common cell wall before the asymmetric cell division. Afterwards, several rounds of anticlinal divisions, in two neighboring cell files, produce a single layer consisting of up to 10 cells, denoted as stage I of LR formation (Malamy and Benfey, 1997;

Dubrovsky et al., 2001; De Rybel et al., 2010). During further development, these cells divide periclinally and form a primordium with two outer and inner layers. Later, cells undergo several anticlinal and periclinal divisions to develop a dome-shaped primordium. Eventually, the newly formed LR emerges through the main root (Malamy and Benfey, 1997).

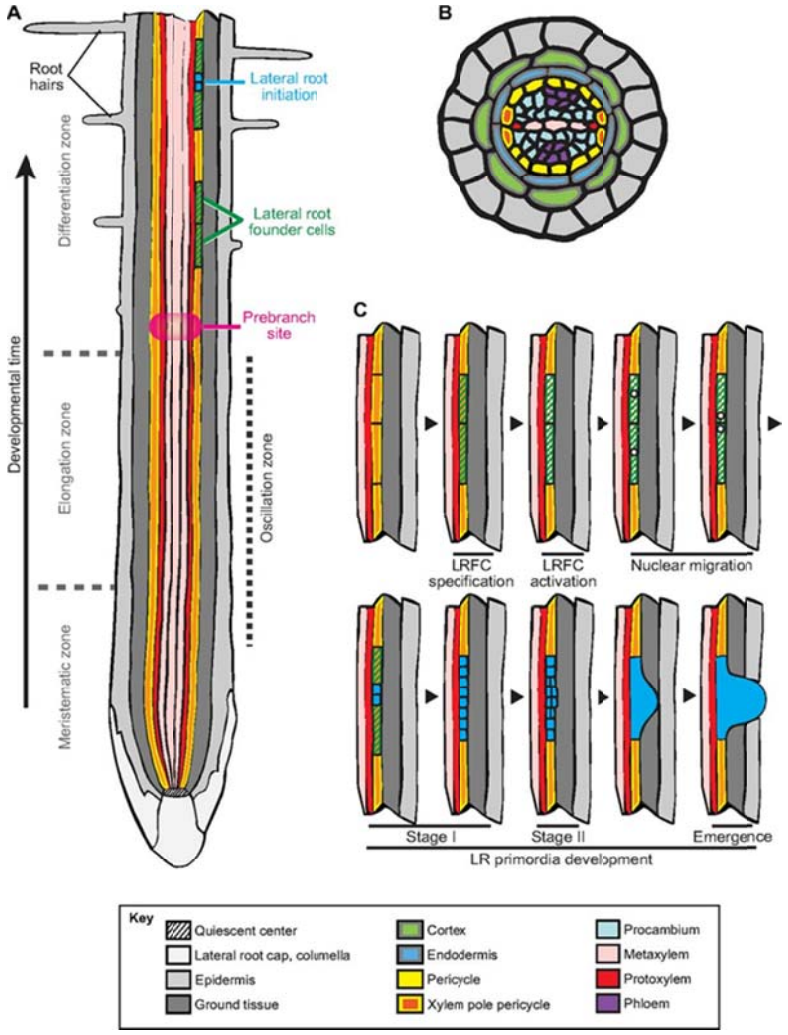


Figure 7: Structure and development of the *Arabidopsis* root. (A) Median longitudinal section depicting developmental time (black arrow) in the longitudinal axis. A prebranch site (magenta) forms after oscillation of gene expression within the oscillation zone (dotted line). Prebranch sites indicate competence to form a lateral root primordium (LRP) in the future. After the competence is established, the xylem pole pericycle (XPP) cells within a prebranch site are predicted to be specified as lateral root founder cells (LRFCs, green hatching). LRP initiate in the differentiation zone through asymmetric cell division of LRFCs, which gives rise to smaller cells (blue). (B) Transverse section. Periodic expression of *DR5:GUS* occurs in the protoxylem; however, because lateral root (LR) initiation occurs in the adjacent XPP cells, signaling between these cell types might be required for LRFC specification. Note that the ground tissue comprises two cell layers: the outermost cortex and the endodermis, which is immediately exterior to the pericycle. (C) Cut-away portion of the median longitudinal section focused on a region where an LR will form. XPP cells are predicted to be sequentially specified as LRFCs (green hatching), then activated to undergo cell division (green/white hatching). LRFC activation results in the coordinated migration of nuclei (white circles) towards the common cell wall in a pair of longitudinally abutted cells. These cells then undergo an asymmetric division, giving rise to smaller cells (blue), to generate a stage-I LRP. The primordium grows through the outer cell layers of the primary root until it emerges from the epidermis. Drawing is not to scale, (taken from Van Norman *et al.* 2013)

II. Peptide hormones in plants

Developmental and environmental responses in multicellular organisms are highly dependent on cell-to-cell communication. Intercellular signaling in plants is particularly important due to the fixed position of cells within rigid cell walls. For many years, the so-called classical phytohormones - auxin, cytokinin, abscisic acid, gibberellin and ethylene - were thought to be the main players of intercellular communication in plants. Later on, several other molecules, including brassinosteroids, jasmonate, salicylic acid and strigolactones, have been added to the list of plant growth regulators (Lindsey et al., 2002; Vanstraelen and Benkova, 2012). However, it is still intriguing how such a relatively small group of molecules can coordinate the huge number of diverse cellular responses. Recent discoveries have made it clear that several other small molecules, including signaling peptides (also called peptide hormones or secreted peptides), small RNAs and transcription factors, are important players in cell-to-cell communication networks (Lindsey et al., 2002; Van Norman et al., 2011; Murphy et al., 2012).

II.1. Structural characteristics and posttranslational modifications of signaling peptides

Peptides are generally defined as small proteins containing 50 or fewer amino acids in their mature form (Murphy et al., 2012). Signaling peptides, once considered to be specific to mammalian systems, are found in plants as well. Systemin was the first plant signaling peptide identified and it was shown to be involved in systemic response to wounding in tomato (Pearce et al., 1991). Following the discovery of systemin, several signaling peptides have been characterized in plants, especially in *Arabidopsis* (Table 1). Nevertheless, considering the myriad of signaling peptides in mammals (Hook et al., 2008), it is likely that many more remain still to be identified.

Based on structural characteristics, signaling peptides are divided into two distinct groups (Figure 1). Group I consists of cysteine-rich peptides, and is divided into two subgroups: peptides requiring intramolecular disulfide bond formation followed by proteolytic processing; and peptides processing multiple intramolecular disulfide bonds that do not undergo processing. Disulfide bonds are crucial for the three-dimensional structure of the protein. Mature cysteine-rich peptides are generally longer than 20 amino acids. Many secreted peptides belong to this class (Silverstein et al., 2007), such as S-locus Cys-rich or S-locus protein 11 (SCR/SP11) (Schopfer et al., 1999; Takayama et al., 2000), STOMAGEN

(Kondo et al., 2010), LURE (Okuda et al., 2009) and EPIDERMAL PATTERNING FACTORs (EPFs) (Hara et al., 2007; Hara et al., 2009). Group II includes cysteine-poor peptides, also called "small posttranslationally modified peptides". In contrast to the aforementioned group, this class is characterized by the small size of the mature peptide (<20 amino acids) due to proteolytic processing, the presence of a C-terminal, often proline-rich, conserved motif that sometimes carries posttranslational modifications, and generally encoded by multiple paralogous genes. Most signaling peptides characterized so far in plants belong to this class. Small posttranslationally modified peptides share a common tripartite structure: (i) a signal peptide at the N-terminal region, (ii) a C-terminal region that is usually conserved among different members of the family and corresponds to the mature peptide and (iii) a variable segment that links the two aforementioned domains (Figure 1). However, in contrast to the majority of the peptides, which are produced by the cleavage of larger precursors, EARLY NODULIN40 (ENOD40) peptides are encoded as such (vandeSande et al., 1996). Interestingly, in several instances, it has been shown that a number of the signaling prepropeptides contain several copies of the C-terminal motif or exhibit a second conserved one in the variable segment (Meng et al., 2012; Roberts et al., 2013). A few other peptides have also been identified that cannot be classified into one of these two groups (Matsubayashi, 2011a, 2011b; Murphy et al., 2012; Aalen, 2013).

Posttranslational modifications contribute to the native conformational structure, and may enhance the receptor-binding capacity. In some studies, lack of posttranslational modification has been shown to reduce the peptide activity (Seitz, 2000; Matsubayashi, 2011a; Shinohara and Matsubayashi, 2013). In plants, the C-terminal region of small posttranslationally modified peptides often carries one or more types of three posttranslational modifications : tyrosine sulfation, which is catalyzed by the plant-specific TYROSYL PROTEIN SULFOTRANSFERASE1 (TPST), proline hydroxylation mediated by PROLYL 4-HYDROXYLASE (P4H) (Myllyharju, 2003) and hydroxyproline arabinosylation. Hydroxyproline arabinosylation is a plant-specific posttranslational modification (Lamport, 1967) and, only recently, three *HYDROXYPROLINE O-ARABINOSYLTRANSFERASE (HPAT)* enzymes have been discovered in *Arabidopsis* (Ogawa-Ohnishi et al., 2013).

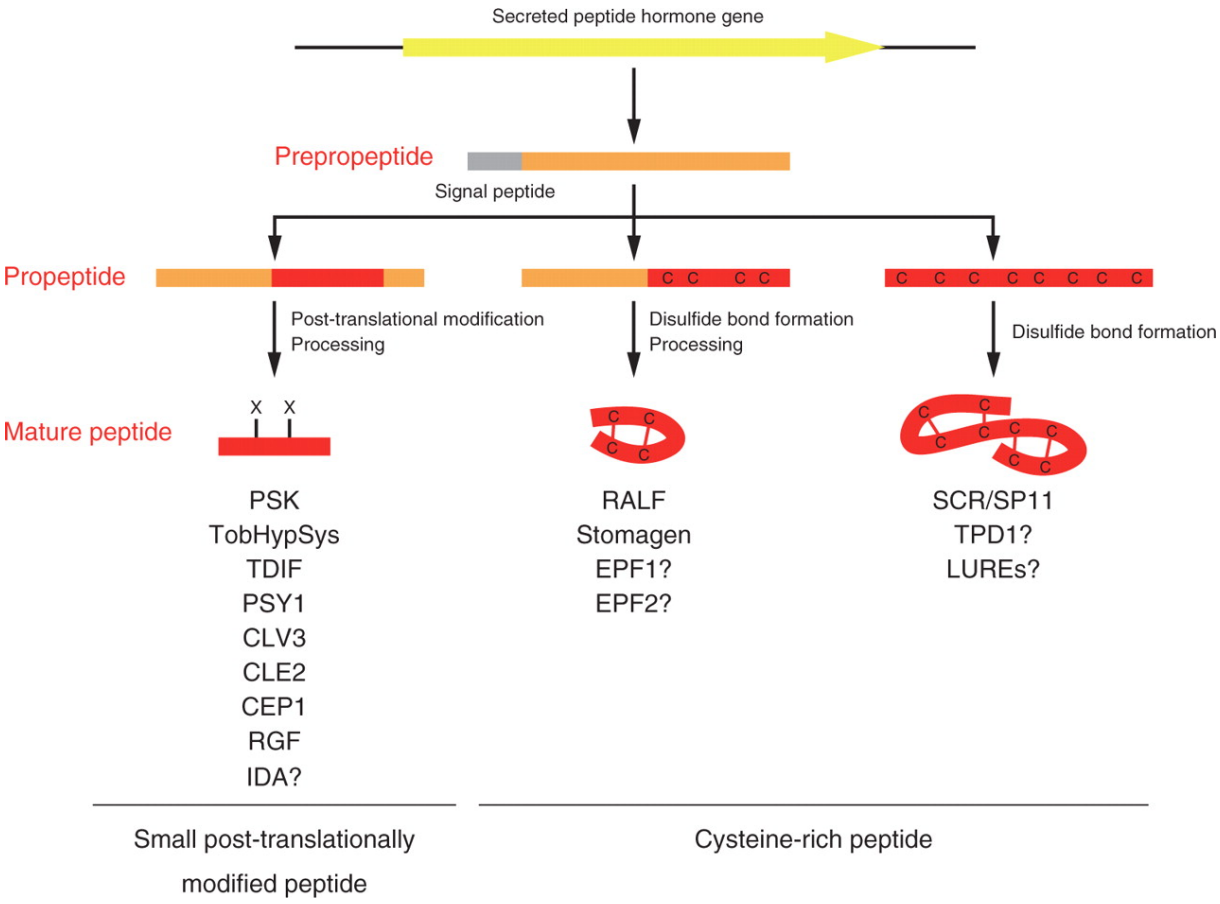


Figure 1: Posttranslational Modifications and Processing of Small Signaling Peptides (taken from Matsubayashi, Y. 2011a).

The mechanisms by which plant-secreted peptides are produced are not well understood, but are thought to be somewhat similar to the maturation process of their mammalian counterparts. In general, signaling peptides are translated as large precursors (so called prepropeptides) followed by one or more posttranslational modifications and/or proteolytic processing. Maturation of signaling peptides is assumed to start by the cotranslational removal of the N-terminal signal peptide by signal peptide peptidases (SPPs) in the endoplasmic reticulum to yield the propeptide (Hook et al., 2008; Matsubayashi, 2011a, 2011b). The cleavage site in the precursor can be predicted with a high degree of accuracy (Petersen et al., 2011). The resulting propeptide is directed through the Golgi apparatus and, together with processing proteases, is loaded into newly formed secretory vesicles. Later on, as the vesicle matures, proteolytic processing occurs with removal of the variable region and release of the mature peptide contained in the C-terminal region (some cysteine-rich peptides do not require the proteolysis step). Finally, mature peptides will be released to the extracellular space and will bind to the cognate receptor(s) in the plasma membrane of

neighboring or, after their transport, more distant cells, where they trigger a signaling pathway to initiate specific physiological responses (Hook et al., 2008).

Membrane-localized receptors function as master switches of complex intracellular signaling. Upon binding of their ligand, receptors autophosphorylate and are activated. Activated receptors will, in turn, phosphorylate downstream targets, which might result in a direct response and/or a phosphorylation cascade, including mitogen-activated protein (MAP)-kinases, that will eventually lead to transcriptional regulation of specific genes (Hook et al., 2008; Matsubayashi, 2011a; Stahl and Simon, 2012).

Table 1: Summary of plant peptide signaling molecules present in *Arabidopsis* thaliana, (Adapted and modified from Wheeler and Irving, 2012).

Propeptide	Gene family	Propeptide size (kDa)	Processed peptide, size	Function	Site of action	Receptor	References
CEP	15	8.5–11.5	CEP1, 14 AA	Inhibits root growth	Lateral root primordia	Unknown	Ohyama et al. (2008) and (Roberts et al., 2013)
CLE (and CLV3)	32	7.8–14.5	CLE or mCLV3, 12–14 AA	Stimulates organogenesis and inhibits meristematic growth; can stimulate vascular development	Floral, shoot, and root meristems; vascular	CLV1, BAM1, CLV2, RPK2	Clark et al. (1995, 1997), (Cock and McCormick, 2001), (Fiers et al., 2006), Kondo et al.(2006), (DeYoung and Clark, 2008), and (Kinoshita et al., 2010)
DVL/RTFL	24	4.6–16.2	Not processed	Polarity dependence, cell proliferation, leaf development	Stem, rosette leaves, pedicles, siliques	Unknown	(Narita et al., 2004), Wenet al. (2004) and (Ikeuchi et al., 2011)
EPF	11	11.5–14.3	Unknown	Promotes epidermal cell division leading to guard cell (stomatal) formation	Epidermis and meristemoid mother cells	TMM, ER, ERL1, ERL2	Hara et al. (2007, 2009) and Hunt and Gray (2009)
IDA and IDL	6	8.4–13	EPIP	Inhibits floral abscission Lateral root primordium development, cell seperation	Abscission zone	HAS, HSL	Butenko et al. (2003) Stenvik et al. (2006) and (Kumpf et al., 2013)
PROPEP	7	9.3–12.3	Pep1, 23 AA	Promotes innate immune responses (a danger signal)	Widespread, leaves	Pep1R	(Huffaker et al., 2006), (Yamaguchi et al., 2006), (Pearce et al., 2008), and (Qi et al., 2010)

PNP	2	13–14	Unknown	Extracellular, cell expansion, water/ion movement, stomatal opening, inhibits ABA-induced stomatal closure	Leaves, mesophyll and guard cells, root stele, stem	Unknown	Gehring et al. (1996), Maryani et al. (2001), Ludidi et al. (2002), Rafudeen et al. (2003), Wang et al. (2007), Gottig et al. (2008), Ruzvidzo et al. (2011), and Wang et al. (2011)
POLARIS	1	4.6	36 AA	Required for root elongation, lateral root formation, leaf vascular patterning	Embryonic root and seedling	Unknown	Casson et al. (2002) and Chilley et al. (2006)
PSK	6	8.7–9.7	PSK-a, 5 AA	Promotes cell proliferation and longevity, root elongation	Widespread, mesophyll cells, roots	PSKR1	Matsubayashi and Sakagami (1996), (Lorbiecke and Sauter, 2002), (Matsubayashi et al., 2002; Matsubayashi et al., 2006), and Kutschmar et al. (2009)
PSY	3	7.9	PSY1, 18 AA	Promotes cellular expansion and proliferation, upregulated by wounding	Mesophyll cells, roots	PSYR1	Amano et al. (2007)
RALF and RALFL	39	7–14	RALF, 25–30 AA	Associated with danger signals, affects growth—inhibits root growth	Widespread in plants	Unknown	Pearce et al. (2001), Silverstein et al. (2007),and (Wu et al., 2007)
GLV/RGF/CLEL	11	13	13-15 AA	Maintain root stem cell Niche, gravitropic response, lateral root and root hair development.	Widespread in plants	Unknown	Matsuzaki et al. (2010) Whitford et al. (2012) Meng et al. (2012) Fernandez et al. (2013)
SCRL	27	9.2–11.5	Not processed	Prevents self-fertilization (but not in A. thaliana)	Pollen	SRK	Schopfer et al. (1999) and (Vanoosthuyse et al., 2001)
TPD	2	19.5	TPD1	Anther development promoting tapetum formation	Anthers	EMS1	(Yang et al., 2003),and (Jia et al., 2008)

II.2. Tools to identify new small secreted peptides

A bottleneck in the functional study of signaling peptides in plant growth and development has been the identification of the encoding genes. Genome sequencing of different plant species has led to the prediction and identification of a number of signaling molecules, but the majority of the secreted peptides are encoded by small genes. The precursors are rarely more than 120 amino acids long and the mature peptide usually less than 20 amino acids. Thus, genes encoding small peptides tend to be absent in genome annotations, being difficult to distinguish them from short, random open reading frames. Additionally, detection of the mature small signaling peptide from crude extracts of plant tissues is not always an easy task given their very low physiological concentration (nanomolar range). Therefore, it is likely that only a few of the total number of secreted peptides are known to date. Identification of new signaling peptide families in plants by various approaches remains a valid goal (Olsen et al., 2002; Lease and Walker, 2006; Fukuda, 2012; Murphy et al., 2012). Three main approaches have been applied individually or in combination with others and can still be used for the identification and isolation of new signaling molecules in plants.

II.2.1. Bioinformatics approaches

Several successful examples of the discovery of signaling peptides by means of *in silico* approaches have been reported. The SignalP algorithm can predict the presence and the location of signal peptide sequences that would target the propeptide to the secretory pathway (Petersen et al., 2011). This algorithm has been used in the detection of several secreted peptides, such as *IDA* and *CEP1* (Butenko et al., 2003; Ohyama et al., 2008). Usually, several criteria are combined to search for new signaling peptides including: (i) the presence of a signal peptide sequence, (ii) the small size of the precursors, (iii) enrichment or lack of cysteine residues in the propeptide, (iv) presence of sites susceptible to posttranslational modification(s), such as tyrosine sulfation (Matsubayashi and Sakagami, 1996; Amano et al., 2007; Matsuzaki et al., 2010; Whitford et al., 2012), proline hydroxylation (Pearce et al., 2001a; Ito et al., 2006; Kondo et al., 2006; Amano et al., 2007; Ohyama et al., 2008; Ohyama et al., 2009; Matsuzaki et al., 2010) and arabinosylation. Application of one or more of the abovementioned features during the *in silico* search led to the identification of some known signal peptide families *in planta*, including *GLV/RGF/CLEL*, *CEP*, *IDA* and some *CLE*

members (Butenko et al., 2003; Ohyama et al., 2008; Matsuzaki et al., 2010; Meng et al., 2012; Murphy et al., 2012; Whitford et al., 2012).

II.2.2. Genetic approaches

Forward and reverse approaches have been used as a tool to identify new signaling peptides. The *clv3* mutant was identified based on an enlarged SAM in an ethyl methanesulfonate–induced population (Clark et al., 1995). Thanks to genetic approaches, several other signaling peptides were identified, such as *IDA* and *TPD1*. Reverse genetics in combination with bioinformatic tools resulted in the identification and characterization of the *GLV/RGF/CLEL* signaling peptide family (Butenko et al., 2003; Ohyama et al., 2008; Matsuzaki et al., 2010; Meng et al., 2012; Murphy et al., 2012; Whitford et al., 2012). Nevertheless, often the lack of T-DNA insertion mutants in genes encoding signaling peptides (because of their small size) and functional redundancy (due to analogous family members) are major problems in applying genetic tools. However, genomic approaches are foreseen to still contribute significantly to the identification of signaling peptides.

Genomic and bioinformatics tools are valuable, but not sufficient in signaling peptide discovery and characterization because (i) there are many un-annotated genes and (ii) these tools may not be particularly useful in non-model and/or unsequenced organisms (Fukuda, 2012; Murphy et al., 2012).

II.2.3. Biochemical approaches

As mentioned above, many characterized signaling peptides harbor posttranslational modifications, such as tyrosine sulfation (Matsubayashi and Sakagami, 1996; Amano et al., 2007; Matsuzaki et al., 2010), proline hydroxylation (Pearce et al., 2001a; Ito et al., 2006; Amano et al., 2007; Ohyama et al., 2008; Ohyama et al., 2009; Matsuzaki et al., 2010) and hydroxyproline arabinosylation (Amano et al., 2007; Ohyama et al., 2009) that are probably necessary for full bioactivity and downstream signaling, but without any specific genomic signature (Seitz, 2000; Matsubayashi, 2011a; Shinohara and Matsubayashi, 2013). Some biochemical methods have been used to test for the presence of bioactive compounds in crude extracts. Therefore, the use of biochemical tools also resulted in the discovery of several plant signaling peptides. For example, systemin, the first plant signaling peptide that had been isolated from extracts of wounded tomato leaves. Alkalization assays led to the identification

of the RAPID ALKALINIZATION FACTOR (RALF) from tobacco and PSK was discovered in conditioned media (Fukuda, 2012; Murphy et al., 2012).

Appreciating the usefulness of each tool alone for the identification of novel peptides, and combining them will certainly enhance the speed and accuracy of the detection of plant signaling peptides.

II.3. Signaling peptides involved in different biological processes

Signaling peptides have been associated with diverse developmental processes in plants, such as apical meristem maintenance (CLAVATA3 (CLV3), GOLVEN/ROOT GROWTH FACTOR/CLE-LIKE [GLV/RGF/CLEL]), defense (systemins), floral organ abscission (INFLORESCENCE DEFICIENT IN ABSCISSION [IDA]), and stomatal patterning (EPF), among others (Wang and Fiers, 2010; Shimada et al., 2011; Fukuda, 2012; Murphy et al., 2012; Stahl and Simon, 2012; Yamada and Sawa, 2013; Fernandez et al., 2013b). To highlight the diverse roles of signaling peptides in various developmental processes, and in particular during root development, some examples of the function of the signaling peptide families are outlined below. As the focus of this thesis is on the GLV/RGF/CLEL family, a special section will be dedicated to these peptides (see below; section III. The GOLVEN signal peptide family).

II.3.1. CLAVATA3/ESR-RELATED (CLE)

The *Arabidopsis* CLAVATA3/ESR-RELATED (CLE) peptide family members were originally identified by the club-shaped fruit phenotype of their corresponding mutants ("clava" Latin for club). Further investigations revealed that the *CLV3* gene is exclusively expressed in a limited number of cells in the shoot apical meristem (SAM) and restrains stem cell activity in the SAM (Clark et al., 1995; Fletcher, 1999; Kondo et al., 2006; Jun et al., 2010). The *clv3* mutants exhibit overproliferation of stem cells, resulting in a meristem almost 1,000 fold larger than that of the wild type. Overproduction of CLV3 causes loss of stem cell activity and differentiation of the shoot meristem (Brand et al., 2000). After binding to the leucine-rich-repeat-receptor like kinase (LRR-RLK) CLV1, the CLV3 peptide promotes the repression of the homeodomain transcription factor *WUSCHEL (WUS)* and, through a negative feedback loop, regulates SAM homeostasis. In this model, WUS binds directly to the *CLV3* promoter and activates its transcription. The CLV3 peptide suppresses *WUS* expression with decreased CLV3 transcription as a result, thereby reducing the number of stem cells (Mayer et al., 1998; Fletcher, 1999; Brand et al., 2000; Schoof et al., 2000; Katsir et al., 2011; Yadav et al., 2011; Murphy et al., 2012).

Additionally, CLV2 and CORYNE (CRN) have been shown to form a receptor complex contributing to SAM maintenance. However, CRN lacks kinase activity and no direct CLV3-CLV2/CRN binding has been demonstrated. Possibly, this receptor complex

relays the CLV3 signal independently from CLV1 (Muller et al., 2008; Nimchuk et al., 2011). Recently, yet another receptor protein, RPK2, was found to relay CLV3 signal (Kinoshita et al., 2010). Interestingly, *clavata*-like phenotype in maize has recently led to the discovery of a downstream component in this signaling pathway (Bommert et al., 2013). *COMPACT PLANT2 (CT2)* encodes the predicted α-subunit (Gα) of a heterotrimeric GTP-binding protein. The maize orthologue of *CLV2*, *FASCIATED EAR2 (FEA2)*, is a receptor without a signaling domain. FEA2 interacts with CT2 to transmit CLAVATA-dependent signals to regulate the SAM maintenance. However, genetic evidence suggested that FEA2 signals through other pathways besides CT2/Gα (Bommert et al., 2013).

A similar molecular mechanism to maintain stem cell activity as described for the SAM (CLV-WUS pathway, see above) has been proposed in the RAM. In the RAM, it is suggested that the binding of the signaling peptide CLE40 to CLV1, activates CLV1 together with the receptor-like kinase ARABIDOPSIS CRINKLY4 (ACR4) This results in the repression of the expression of the transcription factor WUSCHEL-related homeobox 5 (WOX5) that is homologous to WUS. WOX5 controls the communication between the QC and the surrounding stem cells promoting stem cell identity. Plants overexpressing *CLE40* display a stunted primary root growth, lack QC-surrounding stem cells and instead contain a large number of differentiated cells. It is thought that the CLE40-ACR4/CLV1 signaling restrains the expression of WOX5 to the QC, thereby restricting the stem cell identity to the cells in contact with the QC cells. Increased levels of CLE40 will prevent WOX5 from reaching the neighboring cells that will, ultimately, lose their stem cell nature and will differentiate. In *cle40* and *acr4* mutants, the inhibition is lost and the *WOX5* expression domain is expanded resulting in the overproliferation of stem cells (Hobe et al., 2003; Stahl et al., 2009; Stahl et al., 2013)(Figure 2).

The role of plant hormones, transcription factors and microRNAs signals is well established in vasculature development. Recent studies have also highlighted the value of signaling peptides in this context. The TRACHEARY ELEMENT DIFFERENTIATION INHIBITORY FACTOR (TDIF) peptide stimulates procambial cell proliferation in leaf and hypocotyl vasculature and at the same time constrains the differentiation of procambial cells into tracheary elements. TDIF was isolated first from mesophyll cell cultures of *Zinnia elegans*, of which the tracheary element differentiation was arrested. Later investigation led to the isolation of an extracellular 12-amino acid peptide and a homology search revealed that TDIF was identical to the C-terminal motif of CLE41 and CLE44 and highly homologous to

CLE42 and CLE46 in *Arabidopsis* (Fukuda and Komamine, 1980; Kondo et al., 2006). Accordingly, *cle41* and *cle44* mutants show a reduced number of procambial cells and exogenous application of TDIF also increases procambial cell proliferation in *Arabidopsis* hypocotyls (Ito et al., 2006; Hirakawa et al., 2008; Whitford et al., 2008; Hirakawa et al., 2011; Fukuda, 2012). By screening a collection of LRR-RLKs, TDIF RECEPTOR/PHLOEM INTERCALATED WITH XYLEM (*TDR/PXY*) has been identified. *TDR/PXY* is basically expressed in procambial cells, whereas CLE41 is transcribed in neighboring phloem cells, indicating that CLE41 must be secreted toward the xylem where it is perceived by TDR/PXY (Hirakawa et al., 2008; Etchells and Turner, 2010; Hirakawa et al., 2011). Intriguingly, another WUS-related homeobox gene, *WOX4,* is essential for the regulatory function of TDIF in the vascular cell fate, but not for TDIF inhibition of xylem differentiation (Hirakawa et al., 2010; Suer et al., 2011). The signaling peptides CLE9 and CLE10 have also been reported to be expressed in the vasculature and to function in vascular development. *CLE10* overexpression results in severe inhibition of the protoxylem formation in roots. The CLE9/CLE10 peptides are assumed to signal through the CLV2 receptor to inhibit protoxylem formation in roots via cytokinin signaling. Accordingly, protoxylem formation is no longer inhibited by CLE9/CLE10 peptides in *clv2* mutants, indicating a link between the CLE9/CLE10 ligand and the CLV2 receptor in vascular development (Kondo et al., 2011; Fukuda, 2012) (Figure 2).

II.3.2. IDA/IDAL

Abscission is one of the final developmental processes in plants that is coordinated by signaling peptides. Abscission is essential for optimal plant growth, because organs that are not necessary or functional anymore have to be removed through precise and programmed cell separation (Jinn et al., 2000; Stenvik et al., 2006; Stenvik et al., 2008). INFLORESCENCE-DEFICIENT IN ABSCISSION (IDA) and IDA-LIKE (IDL) signaling peptides have been shown to promote organ abscission by stimulating cell separation or by inhibiting cellular repair mechanisms (Butenko et al., 2003; Stenvik et al., 2008). The *IDA* and *IDL* genes are expressed in the cell separation zone that comprises the region where the organs are detached from the plant body (Addicott, 1982). It is believed that they signal through the LRR-RLKs HAESA (HAE) and HAESA-LIKE2 (HSL2) (Stenvik et al., 2008). The *ida* knockout mutants retain floral organs, whereas plants overexpressing the *IDA* or *IDL* genes exhibit premature floral organ abscission, with overproliferation of the abscission zone

and, additionally, ectopic abscission of some organs that normally do not shed in *Arabidopsis* (Jinn et al., 2000; Stenvik et al., 2006; Roberts et al., 2012; Kumpf et al., 2013).

IDA signaling peptides are mostly known for their role in floral organ abscission (Butenko et al., 2003). Recently, a new function has been assigned to the IDA-HAE/HSL2 signaling module, namely a contribution to facilitate the passage of LRP through the main root and to assist in the LR emergence. In roots, *IDA* is strongly and constitutively upregulated, whereas *HAE* and *HSL2* are only transiently upregulated by auxin. Induction of IDA by auxin depends on the auxin influx carrier *LIKE AUX1-3 (LAX3)* and the *AUXIN RESPONSIVE FACTOR7 (ARF7).* Auxin, derived from the tip of the primordium, coordinates cell separation in overlaying LRP tissues through regulation of the IDA peptide. IDA and HAE have been demonstrated to regulate LR emergence via control of pectin degradation (Kumpf et al., 2013).

II.3.3. EPF/EPFL

Stomatal development relies highly on asymmetric cell divisions that need to be precisely coordinated in time and space through cell-to-cell communication networks. Unsurprisingly, signaling peptides have a significant impact on stomatal development (Rowe and Bergmann, 2010; Shimada et al., 2011; Hunt and Gray, 2012). Stomata are usually separated by at least one cell, following the so-called "one-cell spacing-rule" mechanism. Various cysteine-rich signaling peptides that belong to the EPF family, particularly *EPF1*, *EPF2*, *EPF-LIKE6/CHALLAH* (*EPFL6/CHAL*) and *EPFL9/STOMAGEN*, are associated with the regulation of stomatal density and positioning (Hara et al., 2007; Hara et al., 2009; Hunt and Gray, 2009; Abrash and Bergmann, 2010; Abrash et al., 2011; Ohki et al., 2011; Shimada et al., 2011; Lee et al., 2012). Overexpression of *EPF1* and *EPF2* reduces the stomatal density. EPF1 and EPF2 signaling requires the activity of the LRR-receptor-like protein TOO MANY MOUTHS (TMM) and the LRR-RLKs, ERECTA (ER) and ERECTA-LIKE1 (ERL1). EPF1 and EPF2 have been demonstrated to bind to ER and ERL1 receptors and EPF2 to TMM. As TMM lacks an intracellular domain (Shpak et al., 2005), another protein with an extracellular domain is required to mediate the signal transduction. Therefore, TMM interacts directly with the ER receptor and forms a complex that initiates the *EPF* signaling pathway (Hara et al., 2007; Hara et al., 2009; Lee et al., 2012).

Interestingly *EPFL9/STOMAGEN*, another member of the *EPF* family, acts antagonistically to *EPF1* and *EPF2* in stomatal development but also requires *TMM* to

positively control it, indicating that peptide hormones from the same family can have opposite functions through the same receptor *in planta* (Kondo et al., 2010; Sugano et al., 2010). Lastly, the *EPFL6/CHAL* also negatively controls stomatal development in the epidermis of stems and hypocotyls. The ER receptors have been proposed to be involved in the transmission of the *EPFL6/CHAL* signal (Abrash and Bergmann, 2010; Ohki et al., 2011; Shimada et al., 2011).

Recently, a function has been shown for the signaling module EPFL4/EPFL6-ER/ERL in stem elongation (Uchida et al., 2012). In addition, EPFL4 and EPFL6 might be secreted from the endodermis and be perceived by ER/ERL1 in the phloem to regulate vascular development (Uchida and Tasaka, 2013).

II.3.4. PSK and PSY (sulfated peptides)

The only gene responsible for tyrosine sulfation in *Arabidopsis* is *TPST1* and, accordingly, *tpst1* knockout mutants display severe root and shoot phenotypes. *tpst1* mutants have a short root and fail to maintain the root meristem identity, indicating a potentially important function for the sulfated peptides in root development. As expected, several sulfated peptides are reported to be involved in root growth and regulation, namely *GLV/RGF/CLEL*, *PSK-α* and *PLANT PEPTIDE CONTAINING SULFATED TYROSINE1 (PSY1)*. *GLV/RGF/CLEL* peptides mainly coordinate the maintenance of the RAM through regulation of the *PLETHORA (PLT)* function, whereas *PSK-α* and *PSY1* are assumed to control root growth positively by promoting the mature cell size (Matsubayashi et al., 2006; Kutschmar et al., 2009; Matsuzaki et al., 2010; Yamada and Sawa, 2013). Recently a new signaling pathway has been proposed by which the division rate of the QC cells is controlled by *PSK5* (Heyman et al., 2013). In this proposed signaling pathway, the transcription factor ETHYLENE RESPONSIVE FACTOR 115 (ERF115) binds to the promoter region of *PSK5,* thereby activating its expression. The ERF115-PSK5 pathway has been shown to be antagonistally regulated by brassinosteroid signaling and the CELL CYCLE SWITCH 52A2 (CCS52A2) subunit of the ANAPHASE-PROMOTING COMPLEX/CYCLOSOME (APC/C). Brassinosteroids positively regulate ERF115-PSK5, whereas $ACP/C^{CCS52A2}$ reduces the ERF115 abundance through proteolysis (Heyman et al., 2013) (Figure 2).

The sulfated signaling peptide, phytosulfokine-α (PSK-α) positively regulates cell expansion and hypocotyl length in *Arabidopsis*. Through the PSKR1 receptor, the PSK-α signal promotes hypocotyl elongation and protoplast expansion. *tpst1* knockout mutants

displayed severe phenotypes, including shortened hypocotyls, a phenotype that was only partially restored by the supplementation of the PSK-α peptide. This observation implies that, besides PSK-α, likely one or more additional sulfated peptides contribute to the regulation of hypocotyl development (Stuhrwohldt et al., 2011).

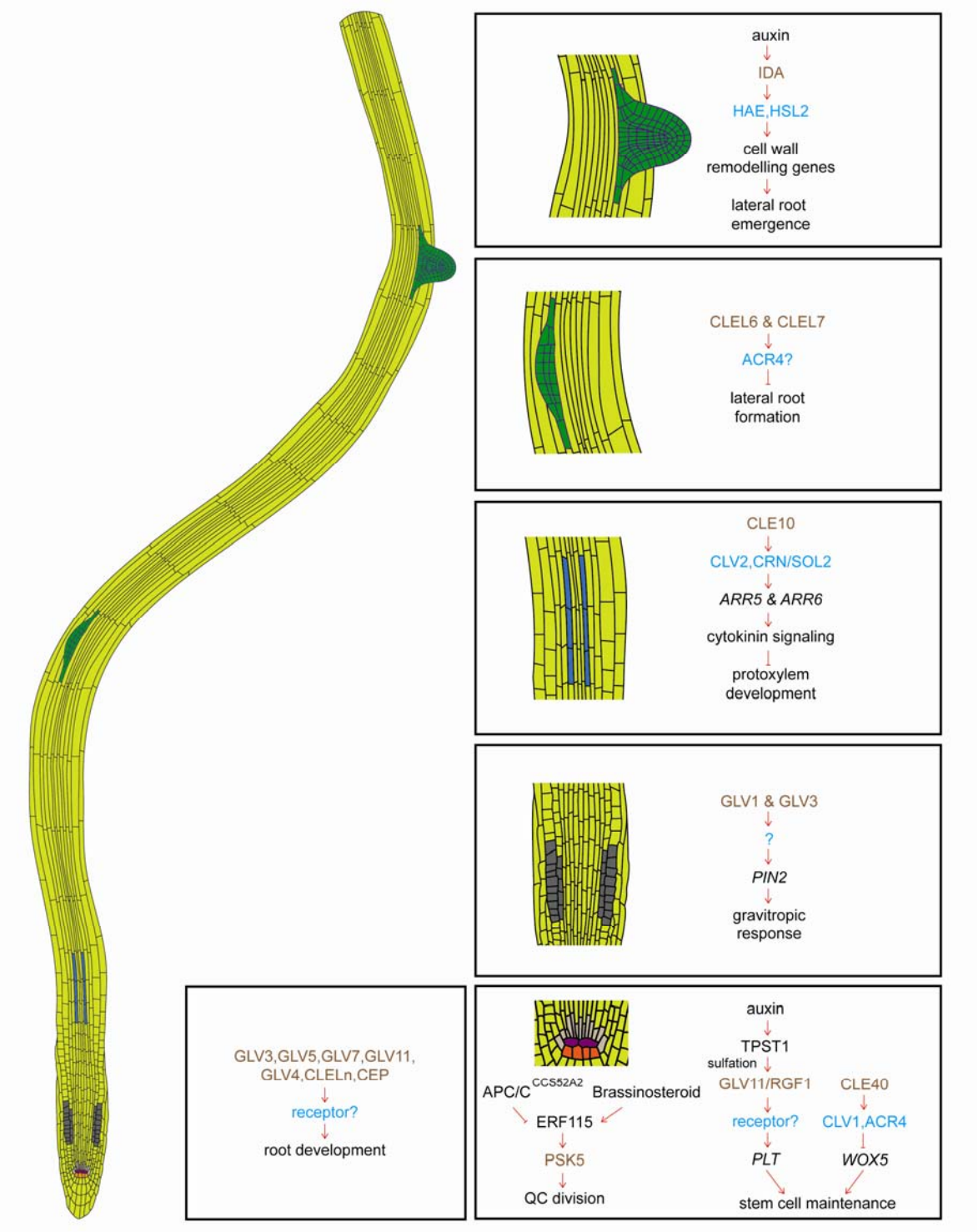


Figure 2: Peptide-mediated regulation of root development**.** CLE, RGF/GLV/CLEL, IDA, and CEP peptides are involved in several aspects of root development including lateral root formation, protoxylem development, stem cell maintenance, gravitropic response, and QC division. Peptides are indicated in brown text and receptors in blue.

III. The GOLVEN signal peptide family

Growth and development are coordinated by an array of cell-to-cell communications (Van Norman et al., 2011; Murphy et al., 2012). Thanks to the completion of the *Arabidopsis* Genome Project (2000), the identification of plant signaling peptides has been accelerated as mentioned above (Lindsey et al., 2002). Three independent *in silico* studies identified a novel signaling peptide family that was designated either *GOLVEN (GLV), ROOT GROWTH FACTOR (RGF),* or *CLE like (CLEL),* with the current redundant nomenclature as a consequence (Table 2). As the three research groups used different sequence homology parameters, not all the family members were initially identified by all groups. The current consensus recognizes 11 members for the *GLV/RGF/CLEL* family in *Arabidopsis* (Matsuzaki et al., 2010; Meng et al., 2012; Whitford et al., 2012). Although *GLV* genes had originally been detected in *Arabidopsis* they are conserved in all higher plants analyzed so far (Whitford et al., 2012). Like most other secreted signaling peptides, they code for posttranslationally modified small peptides. The GLV proteins exhibit a typical tripartite signaling peptide precursor structure (Fernandez et al., 2013b). The sequence of the native secreted peptide has been dissected for four family members, namely GLV1, GLV2, GLV3 and GLV11, which are 14, 15, 18 and 13 amino acids in length, respectively and carry at least two types of the posttranslational modifications in their mature form: tyrosine sulfation and proline hydroxylation. Tyrosine sulfation increases the bioactivity of the mature peptide, but until now hydroxylation of the proline residue has not been associated with any functional role (Matsuzaki et al., 2010; Whitford et al., 2012).

The orchestration of signals within specific developmental frameworks requires the controlled expression and activity of peptides genes. Collectively, the *GLV* genes are transcribed in all plant organs during both vegetative and reproductive stages. Nevertheless, individual transcription patterns are highly specific and are restricted to a few cells and cell types (Figure 3). Nine of the *GLV* genes are active in the root and, based on transcription patterns, *GLV* genes in *Arabidopsis* can be classified in three groups (Fernandez et al., 2013a). Group I encompasses *GLV5*, *GLV6*, *GLV7*, *GLV10* and *GLV11*, which are expressed in the QC and/or columella cells (CCs) and are supposed to be involved in root meristem development. Group II comprises *GLV3*, *GLV6* and *GLV9* that are transcribed in the zone above the QC, mainly in the cortex and endodermis as well as some epidermis and vasculature cells and are assumed to control root growth. Group III is composed of *GLV4* and *GLV8* that are expressed in the region above the root meristem in the epidermis and cortex.

Moreover, ten of these *GLV* genes are expressed during LRP development. Interestingly, the expression pattern of the *GLV* genes follows specific spatiotemporal patterns during LR development. *GLV6*, *GLV5*, *GLV10*, *GLV7* and *GLV11* are transcribed from stage I, II, III and IV, respectively (Malamy and Benfey, 1997), *GLV3* is expressed at stage V and *GLV9* and *GLV2* are expressed after emergence (Fernandez et al., 2013a).

In accordance with the expression profiles in the root tissues, the *GLV* genes are expressed differentially, but are restricted to specific cells or cell types also in the aerial part of the plant: *GLV1*, *GLV2*, *GLV6* and *GLV8* in the shoot and inflorescence. The expression of *GLV1* is irregular in the cotyledon and enhanced at the base of the organ, whereas *GLV2* is transcribed uniformly through the entire cotyledon. Transcription of *GLV1* is also irregular in the leaves and increased at the leaf bases and that of *GLV2* occurs throughout the leaves with enhanced levels in the outer parts. Both genes are expressed in the inflorescence. *GLV1* and *GLV2* are the only members of the *GLV* family that show a promoter activity in the hypocotyl, whereas that of *GLV6* is detected in the vasculature of cotyledons and young leaves as well as in inflorescences. Furthermore, *GLV6* is the only one that is expressed in the SAM. In the cotyledons and leaves, *GLV8* is expressed in a patchy pattern, but its expression can be observed at the base of cotyledon petioles and in stipules. *GLV7* activity has been only detected in pollen. To summarize, *GLV* expression follows very specific and differential patterns, suggesting that the GLV peptides deliver cell-to-cell signals in distinct developmental processes (Fernandez et al., 2013b).

The first loss-of-function phenotype reported for *GLV* genes was a short-root phenotype that has also been observed in the *tpst1* mutant that lacks the activity of GLV/RGF and other peptides. Thus far, TPST is the only identified enzyme that catalyzes tyrosine sulfation in *Arabidopsis* (Komori et al., 2009). Complementation of the other known tyrosine-sulfated signaling peptides PSK and PSY1 could not restore the short-root phenotype, implying the existence of unexplored tyrosine sulfated signaling peptides in *Arabidopsis*. Based on this result, a new signaling peptide family was discovered and designated *RGF* (Matsuzaki et al., 2010). Although single *rgf1/glv11*, *rgf2/glv5*, and *rgf3/glv7* loss-of-function mutants did not exhibit any obvious root phenotypes, the *rgf1 rgf2 rgf3* triple mutant had a short-root phenotype defined by a decreased number of meristematic cortical cells. Complementation of the growth media with a RGF1/GLV11 synthetic peptide could restore the meristem size of both the *tpst* and *rgf1 rgf2 rgf3* mutants. GLV/RGF peptides coordinate the postembryonic maintenance of the root meristem cell niche through defined expression

levels and patterns of the *PLETHORA (PLT)* transcription factors (Matsuzaki et al., 2010) (Figure 2). However, overexpression and exogenous application of all GLV peptides did not have the same effect on root meristem growth, implying that not all members of the family are involved in root meristem maintenance (Matsuzaki et al., 2010; Whitford et al., 2012; Fernandez et al., 2013a).

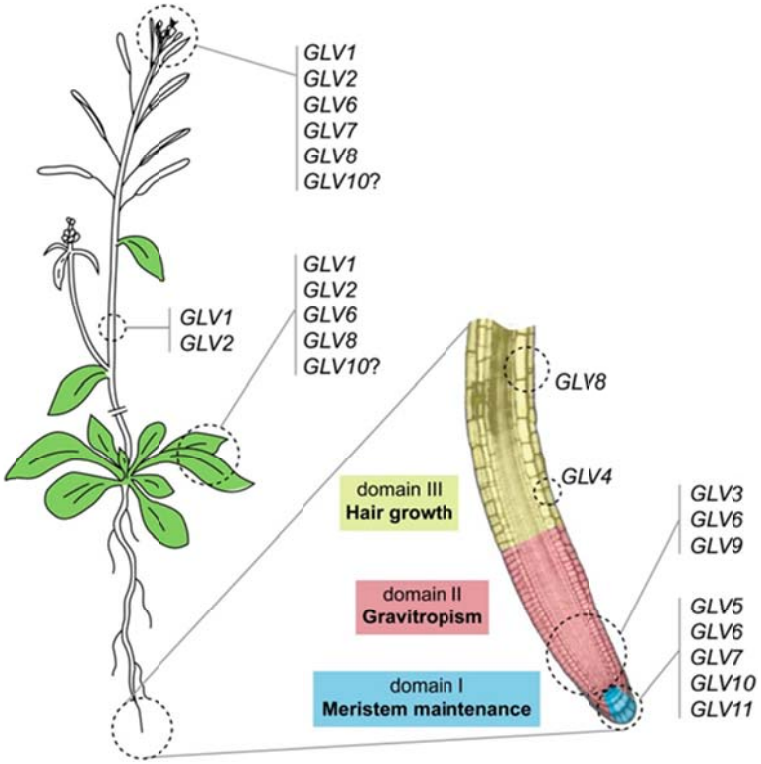


Figure 3: Schematic representation of the *GLV* gene expression domains and functions, (taken from Fernandez *et al.* 2013).

As mentioned above, all *GLV* genes, except *GLV1*, are transcribed during LRP development. Furthermore, overexpression of some *GLV* genes strongly inhibits the successive LR formation and treatment with some GLV synthetic peptides has similar effects on wild-type plants. Despite the lack of information on the LR phenotype in *GLV* loss-of-function mutants, accumulating evidence suggests the involvement of the *GLV* genes in LR initiation and development (Meng et al., 2012; Fernandez et al., 2013a).

A striking curly root phenotype resulting from overexpression of the *GLV* genes revealed a novel function that differed from root growth regulation. Referring to this phenotype, Whitford *et al.* (2012) designated this family *GOLVEN (GLV)*, which means

'waves' in Dutch. This observation suggested a possible role for the *GLV* genes in gravitropic responses. Moreover, rotation of vertically grown *GLV3* gain-of-function mutant seedlings by 90° revealed that the root response to gravistimulation was affected. *GLV3* gain-of-function plants display partially agravitropic roots, whereas the gravitropic response of *amiRglv3* seedlings was enhanced. Addition of the corresponding synthetic peptide mimicked the gain-of-function phenotype. The same altered gravitropic response was observed in hypocotyls of plants overexpressing *GLV1* or *GLV2* genes. The phenotype observed in *GLV*-overexpressing roots resembles that of mutants affected in either auxin response or transport. Further analysis pointed out that *GLV3* is involved in the gravitropic response by influencing PIN-FORMED 2 (PIN2) trafficking. Differential turnover of the auxin efflux carrier PIN2 between the upper and lower side of the gravistimulated root mediates the formation of the lateral auxin gradients by which plants can respond to gravity (Vanneste and Friml, 2009). Overexpression or treatment with GLV peptides hampers PIN2 trafficking, hence preventing the proper formation or maintenance of the required auxin gradient. Most likely, a *GLV* regulatory mechanism increases the robustness of the established auxin gradient, but the molecular mechanism explaining how GLV signal controls the PIN2 trafficking is still unknown (Whitford et al., 2012).

As the expression patterns of *GLV4* and *GLV8* deviate from those of their paralogous counterparts that are not expressed in the lowest part of the root, the corresponding genes are expected to be involved in other root developmental process(es) than root meristem regulation and LR development. The ectopic overexpression of *GLV8* caused the production of root hairs with more complex and irregular shapes, whereas the *glv8* loss-of-function mutant harbors shorter root hairs with a simpler structure than the wild type. *GLV4* silencing also resulted in root hairs that were shorter than those of the wild type, confirming the possible role of these *GLV* genes in root hair development (Fernandez et al., 2013a).

Table 2. Summary of *GLV* expression patterns and mutant phenotypes, (taken from Fernandez et al., 2013).

Gene names	**Promoter reporter**	**Expression pattern**					***gof* phenotypes**				***lof* phenotypes**	**Alteration of PIN2 traffic**
		qRT-PCR	**AtGen Express**	**eFP brower**	**GENEVES TIGATOR**	**Matsuzaki et al., 2010**	**Wavy/curly**	**Increased RAM size**	**Lower LR density**	**Branched root hairs**		
At4g16515 GLV1 RGF6 CLEL6	sh, le, st, fl	sh, le, fl	sh, le, st, fl	fl	sh, le, st, fl	n.d.	+++[a,b,d]	+++[a,b,c,d]	++[a,d]	n.a.	reduced hypocotyl bending[b]	++[b]
At5g64770 GLV2 RGF9 CLEL9	sh, le, st, fl, lr	sh, le, fl	sh, le, st, fl	sh, le, st	sh, le, st, fl	n.d.	+++[a,b]	+++[a,c,d]	+++[a]	n.a.	reduced hypocotyl bending[b]	n.a.
At3g30350 GLV3 RGF4	ro, lr	Ro	ro, fl	ro	ro	n.d.	++[a,b]	+[a,c]	+[a]	+[a]	altered root gravitropism[a,b]	++[a,b]
At3g02240 GLV4 RGF7 CLEL4	ro, lr	Ro	ro,sh, fl	ro	ro, lr	n.d.	+/-[a]	++[a,c]	+[a]	++[a]	shorter root hairs[a]	-
At1g13620 GLV5 RFG2 CLEL1	ro, lr	Ro	n.p.	n.p.	n.p.	ro	++[a]	+++[a,c]	+++[a]	+[a]	short RAM[c,e]	+++[a]
At2g03830 GLV6 RGF8 CLEL2	ro, sh, le, fl, lr	ro, sh, le, fl	ro, fl	ro, sh, le, fl	ro	n.d.	+++[a]	+++[a]	+++[a]	n.a.	n.a.	+[a]
At2g04025 GLV7 RGF3 CLEL3	ro, lr	Ro	ro, fl	ro	ro, lr	ro	++[a]	+++[a,c]	+++[a]	n.a.	short RAM[c,e]	n.a.
At3g02242 GLV8 CLEL5	ro, sh, le, fl, lr	Ro	n.p.	n.p.	n.p.	n.a.	-[a]	+/-[a]	+/-[a]	+++[a]	shorter and fewer root hairs[a]	n.a.

At5g15725 GLV9	ro, lr	Ro	ro	ro, sh	ro, lr	n.a.	++[a]	+[a]	+[a]	n.a.	n.a.	n.a.
At5g51451 GLV10 RGF5 CLEL7	ro, lr	ro, sh, le, fl	n.p.	n.p.	n.p.	n.d.	++[a,d]	++[a,c,d]	+++[a,d]	n.a.	n.a.	n.a.
At5g60810 GLV11 RGF1 CLEL8	ro, lr	Ro	n.p.	n.p.	n.p.	ro	++[a,d]	+++[a,c]	+++[a]	n.a.	short RAM[c,e]	+++[a]

a, Fernandez et al. (2013a); b, Whitford et al. (2012); c, Matsuzaki et al. (2010); d, Meng et al. (2012); e, Observed for *glv5 glv7 glv11* triple mutant (Matsuzaki et al., 2010); gof, gain-of-function; lof, loss-of-function; ro, root; sh, shoot; le, leaves; fl, flowers; st, stem; lr, lateral roots; n.a., not assayed; n.d., not detected; n.p., no probe found in ATH1 GenChip; AtGenExpress Visualization Tool (http://www.weigelworld.org/resources/microarray/AtGenExpress); eFP browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi); GENEVESTIGATOR (https://www.genevestigator.com/gv/plant.jsp).

IV. Serpins: potential regulators of signaling cascades

Signaling peptide precursors are subjected to proteolytic processing to produce active signaling peptides (Hook et al., 2008). However, control of proteolysis is important for plant growth, development, stress and defense responses. The regulation of protease activity is, in turn, necessary to prevent inappropriate processing of the signaling molecules (Van der Hoorn, 2008). The action of proteases is tightly controlled initially at the transcriptional level and later at the protein level by activation of the inactive zymogen (inactive enzyme precursor) and by inhibitors (Turk, 2006) (Figure 4). Based on both gene expression and protein property data, plant serpins are likely to participate in a range of biochemical pathways through regulation of the protease activity (Turk, 2006; Roberts and Hejgaard, 2008). Although the word serpin originally derived from serine protease inhibitor (Carrell and Travis, 1985), serpins are just one of the several serine protease inhibitor families and inhibit other protease families, such as cysteine proteases, as well. Serpins were found in Eukaria, Bacteria, Archaea and viruses and, hence, represent the only protease inhibitors found in all branches of life (Huntington et al., 2000; Patston, 2000; Rawlings et al., 2004; Law et al., 2006; Roberts et al., 2011) .

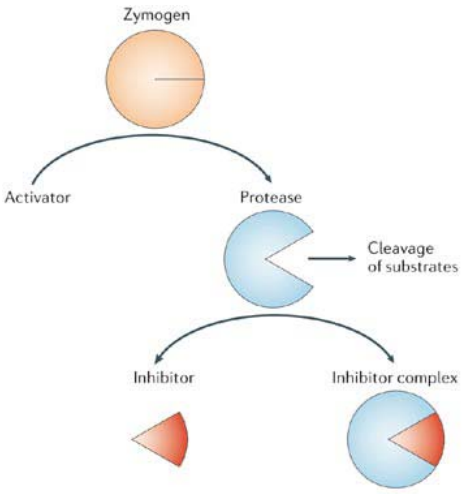


Figure 4: Regulation of protease activity. The fundamental mechanisms governing activity are conserved in most proteases. Latent protease zymogens await an activation signal, which can come from an allosteric activator or another protease. Once active, substrate and inhibitor compete for protease binding, and the outcome is defined by the local inhibitor concentration. The double-headed arrow depicts reversible inhibition (taken from Turk, 2006).

The structure of serpins is highly conserved, but based on their function and regulation individual molecules can form distinct conformations (Silverman and Lomas, 2007). Whereas most serpins control proteolytic cascades, certain do not inhibit enzymes. Inhibitory serpins contain eight to nine α-helices, three β-sheets and a reactive center loop (RCL), which accommodates a specific bait sequence for the target protease(s). The P1 residue of the bait

sequence largely determines the serpin specificity. Proteases initiate the cleavage reaction immediately after the P1 residue. The majority of serpins are irreversible inhibitors, but a small number of serpins inhibit specific proteases in a reversible manner. When serpins interact with a protease, a covalent serpin-protease complex is formed. Furthermore, massive and irreversible conformational changes that take place in the serpin inhibit both the serpin and the protease. (Huntington et al., 2000; Ye and Goldsmith, 2001; Zhou et al., 2001; Roberts and Hejgaard, 2008; Fluhr et al., 2012) (Figure 5).

There is no consensus regarding the number of the genes coding for serpins in *Arabidopsis*. PSI-BLAST analysis has demonstrated that in the *Arabidopsis* genome only six out of ~25 serpin genes and pseudogenes probably encode functional serpins, of which one was predicted to have a non-inhibitory function and five an inhibitory one (Roberts and Hejgaard, 2008).

Owing to their medical importance, extensive progress has been made in understanding the function of mammalian serpins. Unlike the mammalian serpins, only two specific target proteases have been reported for plant serpins so far. Plant serpins have probably functions that differ from their mammalian counterparts, because most of the mammalian serpins are involved in mammal-specific processes, such as blood coagulation, that are absent in plants (Silverman et al., 2010; Roberts et al., 2011). By means of a yeast two-hybrid screen, a serpin homolog was identified in *Arabidopsis* as an endogenous inhibitor of metacaspase 9 (AtMC9), which is the first target for serpins characterized. Recently, the vacuolar protease RESPONSIVE-TO-DESICCATION-21 (RD21) has been reported to be inhibited by AtSERPIN1 (AtSRP1) (Vercammen et al., 2006; Lampl et al., 2010; Lampl et al., 2013). Furthermore, AtSRP2 and AtSRP3 have been suggested to have a regulatory function in DNA damage response (Ahn et al., 2009). As AtSRP1 has been shown to inhibit proteases from all pests *in vitro* (Alvarez-Alfageme et al., 2011), its use has been proposed in pest control. Although serpins are predicted to be most probably involved in a broad range of biochemical processes in plants, knowledge about their function remains very poor. Although serpins are good candidates for propeptide processing regulation, no function has been reported so far for serpins or other protein inhibitors during signaling peptide processing.

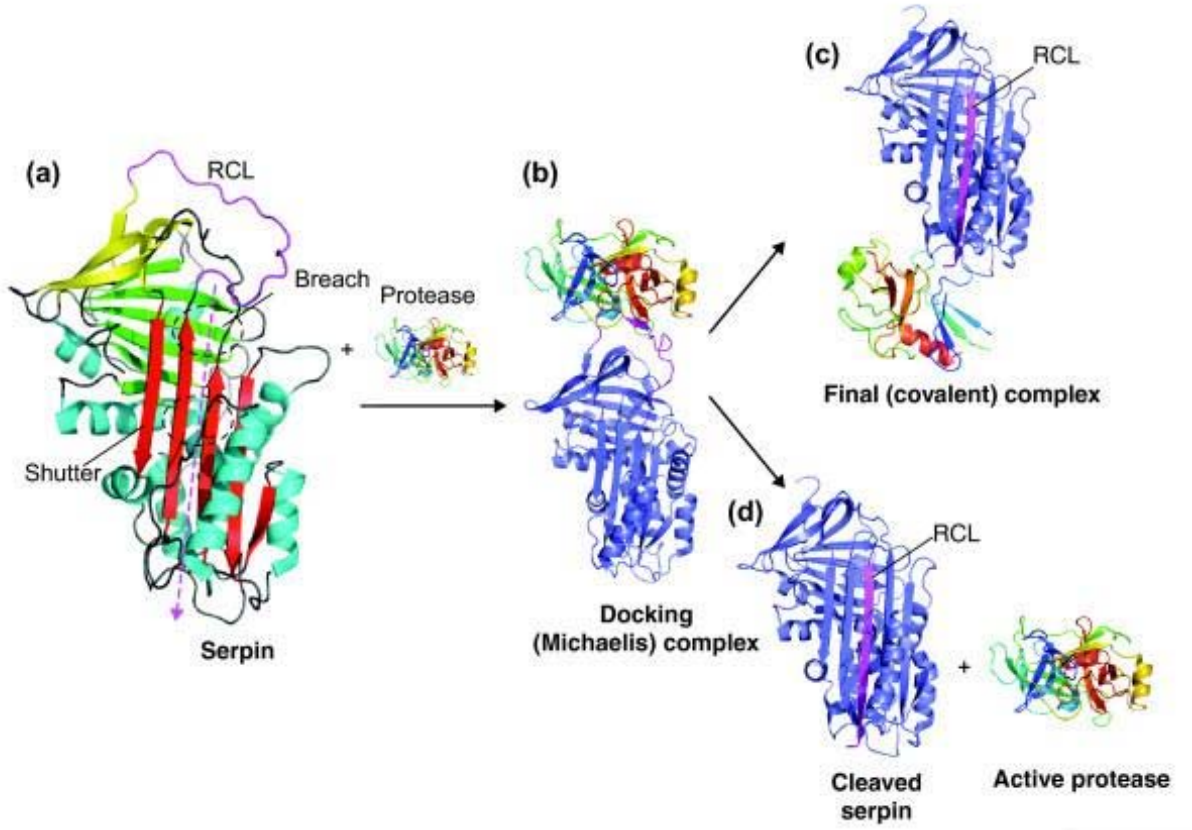


Figure 5: The structure and mechanism of inhibitory serpins. (a) The structure of native SRP1 (Protein Data Bank [PDB] code 1QLP). The A sheet is in red, the B sheet in green and the C sheet in yellow; helices (hA-hI) are in blue. The reactive center loop (RCL) at the top of the molecule is in magenta. The position of the breach and the shutter are labeled and the path of the RCL insertion indicated (magenta dashed line). Both of these regions contain several highly conserved residues, many of which are mutated in various serpinopathies. (b) The Michaelis or docking complex between SRP1 and inactive trypsin (PDB code 1OPH), with the protease (multicolors) docked onto the RCL (magenta). Upon docking with an active protease (b), two possible pathways are apparent. (c) The final serpin enzyme complex (PDB code 1EZX [12]). The serpin has undergone the S-to-R transition and the protease hangs distorted at the base of the molecule. (d) The structure of the cleaved SRP1 is shown (PDB code 7API) with the RCL (magenta) forming the fourth strand of the A sheet. The result of the serpin substrate-like behavior can be seen where the protease has escaped the conformational trap, leaving active protease and inactive, cleaved serpin. Certain serpin mutations, particularly nonconservative substitutions within the hinge region of the RCL, result in a substrate-like, rather than an inhibitory behavior (taken from Law *et al*. 2006).

References

Aalen RB (2013) Maturing peptide open for communications. J Exp Bot **64**: 5231-5235

Abrash EB, Davies KA, Bergmann DC (2011) Generation of signaling specificity in *Arabidopsis* by spatially restricted buffering of ligand--receptor interactions. Plant Cell **23**: 2864-2879

Abrashn EB, Bergmann DC (2010) Regional specification of stomatal production by the putative ligand CHALLAH. Development **137**: 447-455

Addicott FT (1982) *Abscission*. Berkeley, University of California Press

Ahn J-W, Atwell BJ, Roberts TH (2009) Serpin genes *AtSRP2* and *AtSRP3* are required for normal growth sensitivity to a DNA alkylating agent in *Arabidopsis*. BMC Plant Biol **9**: 52

Alvarez-Alfageme F, Maharramov J, Carrillo L, Vandenabeele S, Vercammen D, Van Breusegem F, Smagghe G (2011) Potential use of a serpin from *Arabidopsis* for pest control. PLoS ONE **6**: e20278

Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. Proc Natl Acad Sci USA **104**: 18333-18338

Baluška F, Barlow PW, Baskin TI, Chen R, Feldman L, Forde BG, Geisler M, Jernstedt J, Menzel D, Muday GK, Murphy A, Šamaj J, Volkmann D (2005) What is apical and what is basal in plant root development? Trends Plant Sci **10**: 409-411

Baluška F, Mancuso S, Volkmann D, Barlow PW (2010) Root apex transition zone: a signalling--response nexus in the root. Trends Plant Sci **15**: 402-408

Barlow PW (2002) Cellular patterning in root meristems: its origins and significance. In *Plant Roots, The Hidden Half*, 3rd ed, Y Waisel, A Eshel, U Kafkafi (Eds). New York, Marcel Dekker, pp 49-82

Beeckman T, Burssens S, Inzé D (2001) The peri-*cell*-cycle in *Arabidopsis*. J Exp Bot **52**: 403-411

Bennett T, Scheres B (2010) Root Development--Two meristems for the price of one? Curr Topics Dev Biol **91**: 67-102

Bommert P, Je BI, Goldshmidt A, Jackson D (2013) The maize Gα gene *COMPACT PLANT2* functions in CLAVATA signalling to control shoot meristem size. Nature **502**: 555-558

Brand U, Fletcher JC, Hobe M, Meyerowitz EM, Simon R (2000) Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by *CLV3* activity. Science **289**: 617-619

Butenko MA, Patterson SE, Grini PE, Stenvik G-E, Amundsen SS, Mandal A, Aalen RB (2003) *INFLORESCENCE DEFICIENT IN ABSCISSION* controles floral organ abscission in Arabidopsis and identifies a novel family of putative ligands in plants. Plant Cell **15**: 2296-2307

Carrell R, Travis J (1985) α_1-Antitrypsin and the serpins: variation and countervariation. Trends Biochem Sci **10**: 20-24

Charlton WA (1996) Lateral root initiation. In *Plant Roots: the Hidden Half*, Y Waisel, A Eshel, U Kafkafi (Eds), 2nd ed. New York, Marcel Dekker, pp 149-173

Clark SE, Running MP, Meyerowitz EM (1995) *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. Development **121**: 2057-2067

Clowes FAL (1961) *Apical Meristems*. Oxford, Blackwell Scientific Publications

Cock JM, McCormick S (2001) A large family of genes that share homology with *CLAVATA3*. Plant Physiol **126**: 939-942

De Rybel B, Vassileva V, Parizot B, Demeulenaere M, Grunewald W, Audenaert D, Van Campenhout J, Overvoorde P, Jansen L, Vanneste S, Möller B, Wilson M, Holman T, Van Isterdael G, Brunoud G, Vuylsteke M, Vernoux T, De Veylder L, Inzé D, Weijers D, Bennett MJ, Beeckman T (2010) A novel Aux/IAA28 signaling cascade activates GATA23-dependent specification of lateral root founder cell identity. Curr Biol **20**: 1697-1706

De Smet I, Tetsumura T, De Rybel B, Frey dit Frei N, Laplaze L, Casimiro I, Swarup R, Naudts M, Vanneste S, Audenaert D, Inzé D, Bennett MJ, Beeckman T (2007) Auxin-dependent regulation of lateral root positioning in the basal meristem of *Arabidopsis*. Development **134**: 681-690

DeYoung BJ, Clark SE (2008) BAM receptors regulate stem cell specification and organ development through complex interactions with CLAVATA signaling. Genetics **180**: 895-904

Dolan L, Janmaat K, Willemsen V, Linstead P, Poethig S, Roberts K, Scheres B (1993) Cellular organisation of the *Arabidopsis thaliana* root. Development **119**: 71-84

Dubrovsky JG, Rost TL (2012) Pericycle. In *Encyclopedia of Life Sciences* (ELS). Chichester, John Wiley & Sons

Dubrovsky JG, Rost TL, Colón-Carmona A, Doerner P (2001) Early primordium morphogenesis during lateral root initiation in *Arabidopsis thaliana*. Planta **214**: 30-36

Etchells JP, Turner SR (2010) The PXY-CLE41 receptor ligand pair defines a multifunctional pathway that controls the rate and orientation of vascular cell division. Development **137**: 767-774

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, Hilson P (2013) Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol **161**: 954-970

Fernandez A, Hilson P, Beeckman T (2013) GOLVEN peptides as important regulatory signalling molecules of plant development. J Exp Bot **64**: 5263-5268

Fiers M, Golemiec E, van der Schors R, van der Geest L, Li KW, Stiekema WJ, Liu C-M (2006) The CLAVATA3/ESR motif of CLAVATA3 is functionally independent from the nonconserved flanking sequences. Plant Physiol **141**: 1284-1292

Fitter A (2002) Characteristic and functions of root systems. In *Plant Roots, The Hidden Half*, 3rd ed, Y Waisel, A Eshel, U Kafkafi (Eds). New York, Marcel Dekker, pp 15-32

Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM (1999) Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science **283**: 1911-1914

Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM (1999) Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science **283**: 1911-1914

Fluhr R, Lampl N, Roberts TH (2012) Serpin protease inhibitors in plant biology. Physiol Plant **145**: 95-102

Fukuda H (2012) Peptide regulating plant vascular development. In *Plant Signaling Peptides*, Signaling and Communication in Plants Series, Vol 16, HR Irving, C Gehring (Eds). Berlin, Springer, pp 59-75

Fukuda H, Komamine A (1980) Establishment of an experimental system for the study of tracheary element differentiation from single cells isolated from the mesophyll of *Zinnia elegans*. Plant Physiol **65**: 57-60

Gallagher KL (2013) Cellular patterning of the root meristem: genes and signals. In *Plant Roots: The Hidden Half*, 4th ed, A Eshel, T Beeckman (Eds). Boca Raton, FL, CRC Press, pp 3.1-3.26

Guyomarc'h S, Lucas M, Laplaze L (2010) Lateral/secondary roots. In *Encyclopedia of Life Sciences* (eLS). Chichester, John Wiley & Sons (doi: 10.1002/9780470015902.a0002060.pub2)

Hara K, Kajita R, Torii KU, Bergmann DC, Kakimoto T (2007) The secretory peptide gene *EPF1* enforces the stomatal one-cell-spacing rule. Genes Dev **21**: 1720-1725

Hara K, Yokoo T, Kajita R, Onishi T, Yahata S, Peterson KM, Torii KU, Kakimoto T (2009) Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in Arabidopsis leaves. Plant Cell Physiol **50**: 1019-1031

Heyman J, Cools T, Vandenbussche F, Heyndrickx KS, Van Leene J, Vercauteren I, Vanderauwera S, Vandepoele K, De Jaeger G, Van Der Straeten D, De Veylder L (2013) ERF115 controls root quiescent center cell division and stem cell replenishment. Science **342**: 860-863

Hirakawa Y, Kondo Y, Fukuda H (2010) TDIF peptide signaling regulates vascular stem cell proliferation via the *WOX4* homeobox gene in *Arabidopsis*. Plant Cell **22**: 2618-2629

Hirakawa Y, Kondo Y, Fukuda H (2011) Establishment and maintenance of vascular cell communities through local signaling. Curr Opin Plant Biol **14**: 17-23

Hirakawa Y, Shinohara H, Kondo Y, Inoue A, Nakanomyo I, Ogawa M, Sawa S, Ohashi-Ito K, Matsubayashi Y, Fukuda H (2008) Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. Proc Natl Acad Sci USA **105**: 15208-15213

Hobe M, Müller R, Grünewald M, Brand U, Simon R (2003) Loss of CLE40, a protein functionally equivalent to the stem cell restricting signal CLV3, enhances root waving in *Arabidopsis*. Dev Genes Evol **213**: 371-381

Hook V, Funkelstein L, Lu D, Bark S, Wegrzyn J, Hwang S-R (2008) Proteases for processing proneuropeptides into peptide neurotransmitters and hormones. Annu Rev Pharmacol Toxicol **48**: 393-423

Huffaker A, Pearce G, Ryan CA (2006) An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. Proc Natl Acad Sci USA **103**: 10098-10103

Hunt L, Gray JE (2009) The signaling peptide EPF2 controls asymmetric cell divisions during stomatal development. Curr Biol **19**: 864-869

Hunt L, Gray JE (2012) Peptides modulating development of specialized cells. In *Plant Signaling Peptides*, Signaling and Communication in Plants Series, Vol 16, HR Irving, C Gehring (Eds). Berlin, Springer, pp 93-106

Huntington JA, Read RJ, Carrell RW (2000) Structure of a serpin--protease complex shows inhibition by deformation. Nature **407**: 923-926

Ikeuchi M, Yamaguchi T, Kazama T, Ito T, Horiguchi G, Tsukaya H (2011) ROTUNDIFOLIA4 regulates cell proliferation along the body axis in Arabidopsis shoot. Plant Cell Physiol **52**: 59-69

Ito Y, Nakanomyo I, Motose H, Iwamoto K, Sawa S, Dohmae N, Fukuda H (2006) Dodeca-CLE peptides as suppressors of plant stem cell differentiation. Science, 313, 842-845

Jansen L, Demeulenaere M, Beeckman T (2013) Lateral root development. In *Plant Roots: The Hidden Half*, 4th ed, A Eshel, Beeckman (Eds). Boca Raton, FL, CRC Press, pp 6.1-6.13

Jia G, Liu X, Owen HA, Zhao D (2008) Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. Proc Natl Acad Sci USA **105**: 2220-2225

Jinn T-L, Stone JM, Walker JC (2000) *HAESA*, an *Arabidopsis* leucine-rich repeat receptor kinase, controls floral organ abscission. Genes Dev **14**: 108-117

Jun J, Fiume E, Roeder AHK, Meng L, Sharma VK, Osmont, KS Baker, C Ha, CM Meyerowitz, EM Feldman, LJ Fletcher JC (2010) Comprehensive analysis of *CLE* polypeptide signaling gene expression and overexpression activity in Arabidopsis. Plant Physiol **154**: 1721-1736

Katsir L, Davies KA, Bergmann DC, Laux T (2011) Peptide signaling in plant development. Curr Biol **21**: R356-R364

Kinoshita A, Betsuyaku S, Osakabe Y, Mizuno S, Nagawa S, Stahl Y, Simon R, Yamaguchi-Shinozaki K, Fukuda H, Sawa S (2010) RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in *Arabidopsis*. Development **137**: 3911-3920 [Err. 137, 4327]

Komori R, Amano Y, Ogawa-Ohnishi M, Matsubayashi Y (2009) Identification of tyrosylprotein sulfotransferase in *Arabidopsis*. Proc Natl Acad Sci USA **106**: 15067-15072

Kondo T, Kajita R, Miyazaki A, Hokoyama M, Nakamura-Miura T, Mizuno S, Masuda Y, Irie K, Tanaka Y, Takada S, Kakimoto T, Sakagami Y (2010) Stomatal density is controlled by a mesophyll-derived signaling molecule. Plant Cell Physiol **51**: 1-8

Kondo T, Sawa S, Kinoshita A, Mizuno S, Kakimoto T, Fukuda H, Sakagami Y (2006) A plant peptide encoded by *CLV3* identified by in situ MALDI-TOF MS analysis. Science **313**: 845-848

Kondo Y, Hirakawa Y, Kieber JJ, Fukuda H (2011) CLE peptides can negatively regulate protoxylem vessel formation via cytokinin signaling. Plant Cell Physiol **52**: 37-48

Kumpf RP, Shi C-L, Larrieu A, Stø IM, Butenko MA, Péret B, Riiser ES, Bennett MJ, Aalen RB (2013) Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence. Proc Natl Acad Sci USA **110**: 5235-5240

Kutschmar A, Rzewuski G, Stührwohldt N, Beemster GTS, Inzé D, Sauter M (2009) PSK-α promotes root growth in Arabidopsis. New Phytol **181**: 820-831

Kutschmar A, Rzewuski G, Stührwohldt N, Beemster GTS, Inzé D, Sauter M (2009) PDK-α promotes root growth in Arabidopsis. New Phytol **181**: 820-831

Lampl N, Alkan N, Davydov O, Fluhr R (2013) Set-point control of RD21 protease activity by AtSerpin1 controls cell death in Arabidopsis. Plant J **74**: 498-510

Lampl N, Budai-Hadrian O, Davydov O, Joss TV, Harrop SJ, Curmi PMG, Roberts TH, Fluhr R (2010) *Arabidopsis* AtSerpin1, crystal structure and *in vivo* interaction with its target protease RESPONSIVE TO DESICCATION-21 (RD21). J Biol Chem **285**: 13550-13560

Lamport DTA (1967) Hydroxyproline-*O*-glycosidic linkage of the plant cell wall glycoprotein extensin. Nature **216**: 1322-1324

Lavenus J, Goh T, Roberts I, Guyomarc'h S, Lucas M, De Smet I, Fukaki H, Beeckman T, Bennett M, Laplaze L (2013) Lateral root development in *Arabidopsis*: fifty shades of auxin. Trends Plant Sci **18**: 450-458

Law RHP, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI, Whisstock JC (2006) An overview of the serpin superfamily. Genome Biol **7**: 216

Lease KA, Walker JC (2006) The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol **142**: 831-838

Lee JS, Kuroha T, Hnilova M, Khatayevich D, Kanaoka MM, McAbee JM, Sarikaya M, Tamerler C, Torii KU (2012) Direct interaction of ligand-receptor pairs specifying stomatal patterning. Genes Dev **26**: 126-136

Lindsey K, Casson S, Chilley P (2002) Peptides: new signalling molecules in plants. Trends Plant Sci **7**: 78-83

Lorbiecke R, Sauter M (2002) Comparative analysis of PSK peptide growth factor precursor homologs. Plant Sci **163**: 321-332

Malamy JE (2005) Intrinsic and environmental response pathways that regulate root system architecture. Plant Cell Environ **28**: 67-77

Malamy JE, Benfey PN (1997) Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. Development **124**: 33-44

Matsubayashi Y (2011a) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol **52**: 5-13

Matsubayashi Y (2011b) Small post-translationally modified peptide signals in Arabidopsis. Arabidopsis Book 9, e0150.

Matsubayashi Y, Sakagami Y (2006) Peptide hormones in plants. Annu Rev Plant Biol **57**: 649-674

Matsubayashi Y, Ogawa M, Morita A, Sakagami Y (2002) An LRR receptor kinase involved in perception of a peptide plant hormone, phytosulfokine. Science **296**: 1470-1472

Matsubayashi Y, Sakagami Y (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. Proc Natl Acad Sci USA **93**: 7623-7627

Matsuzaki Y, Ogawa-Ohnishi M, Mori A, Matsubayashi Y (2010) Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. Science **329**: 1065-1067

Mayer KFX, Schoof H, Haecker A, Lenhard M, Jürgens G, Laux T (1998) Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. Cell **95**: 805-815

Meng L, Buchanan BB, Feldman LJ, Luan S (2012) CLE-like (CLEL) peptides control the pattern of root growth and lateral root development in *Arabidopsis*. Proc Natl Acad Sci USA **109**: 1760-1765

Moreno-Risueno MA, Van Norman JM, Moreno A, Zhang J, Ahnert SE, Benfey PN (2010) Oscillating gene expression determines competence for periodic *Arabidopsis* root branching. Science **329**: 1306-1311

Müller R, Bleckmann A, Simon R (2008) The receptor kinase CORYNE of *Arabidopsis* transmits the stem cell-limiting signal CLAVATA3 independently of CLAVATA1. Plant Cell **20**: 934-946

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in *Arabidopsis* development: how cells communicate over a short distance. Plant Cell **24**: 3198-3217

Myllyharju J (2003) Prolyl 4-hydroxylases the key enzymes of collagen biosynthesis. Matrix Biol **22**: 15-24

Narita NN, Moore S, Horiguchi G, Kubo M, Demura T, Fukuda H, Goodrich J, Tsukaya H (2004) Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*. Plant J **38**: 699-713

Nimchuk ZL, Tarr PT, Ohno C, Qu X, Meyerowitz EM (2011) Plant stem cell signaling involves ligand-dependent trafficking of the CLAVATA1 receptor kinase. Curr Biol **21**: 345-352

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y (2013) Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. Nat Chem Biol **9**: 726-730

Ohki S, Takeuchi M, Mori M (2011) The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. Nat Commun **2**: 512

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis. Plant J **55**: 152-160

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening followed by LC-MS-based structure analysis. Plant J **55**: 152-160

Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y (2009) A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*. Nat Chem Biol **5**: 578-580

Okuda S, Tsutsui H, Shiina K, Sprunck S, Takeuchi H, Yui R, Kasahara RD, Hamamura Y, Mizukami A, Susaki D, Kawano N, Sakakibara T, Namiki S, Itoh K, Otsuka K, Matsuzaki M, Nozaki H, Kuroiwa T, Nakano A, Kanaoka MM, Dresselhaus T, Sasaki N, Higashiyama T (2009) Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature **458**: 357-361

Olsen AN, Mundy J, Skriver K (2002) Peptomics, identification of novel cationic *Arabidopsis* peptides with conserved sequence motifs. In Silico biol **2**: 441-451

Osmont KS, Sibout R, Hardtke CS (2007) Hidden branches: developments in root system architecture. Annu Rev Plant Biol **58**: 93-113

Patston PA (2000) Serpins and other serine protease inhibitors. Immunol Today **21**: 354-354

Pearce G, Moura DS, Stratmann J, Ryan CA (2001a) Production of multiple plant hormones from a single polyprotein precursor. Nature **411**: 817-820

Pearce G, Strydom D, Johnson S, Ryan CA (1991) A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science **253**: 895-898

Pearce G, Yamaguchi Y, Munske G, Ryan CA (2008) Structure--activity studies of AtPep1, a plant peptide signal involved in the innate immune response. Peptides **29**: 2083-2089

Perilli S, Di Mambro R, Sabatini S (2012) Growth and development of the root apical meristem. Curr Opin Plant Biol **15**: 17-23

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8**: 785-786

Qi Z, Verma R, Gehring C, Yamaguchi Y, Zhao Y, Ryan CA, Berkowitz GA (2010) Ca^{2+} signaling by plant *Arabidopsis thaliana* Pep peptides depends on AtPepR1, a receptor with guanylyl cyclase activity, and cGMP-activated Ca^{2+} channels. Proc Natl Acad Sci USA **107**: 21193-21198

Rawlings ND, Tolle DP, Barrett AJ (2004) MEROPS: the peptidase database. Nucleic Acids Res **32**: D160-D164

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in *Arabidopsis* shoot development. J Exp Bot **64**: 5371-5381

Roberts IN, Caputo C, Criado MV, Funk C (2012) Senescence-associated proteases in plants. Physiol Plant **145**: 130-139

Roberts TH, Ahn J-W, Lampl N, Fluhr R (2011) Plants and the study of serpin biology. Methods Enzymol **499**: 347-366

Roberts TH, Hejgaard J (2008) Serpins in plants and green algae. Funct Integr Genomics **8**: 1-27

Rowe MH, Bergmann DC (2010) Complex signals for simple cells: the expanding ranks of signals and receptors guiding stomatal development. Curr Opin Plant Biol 13, 548-555

Schoof H, Lenhard M, Haecker A, Mayer KFX, Jürgens G, Laux T (2000) The stem cell population of *Arabidopsis* shoot meristem is maintained by a regulatory loop between the *CLAVATA* and *WUSCHEL* genes. Cell **100**: 635-644

Schopfer CR, Nasrallah ME, Nasrallah JB (1999) The male determinant of self-incompatibility in *Brassica*. Science **286**: 1697-1700

Seitz O (2000) Glycopeptide synthesis and the effects of glycosylation on protein structure and activity. ChemBioChem **1**: 214-246

Shimada T, Sugano SS, Hara-Nishimura I (2011) Positive and negative peptide signals control stomatal density. Cell Mol Life Sci **68**: 2081-2088

Shinohara H, Matsubayashi Y (2013) Chemical synthesis of Arabidopsis CLV3 glycopeptide reveals the impact of hydroxyproline arabinosylation on peptide conformation and activity. Plant Cell Physiol **54**: 369-374

Shpak ED, Messmer McAbee J, Pillitteri LJ, Torii KU (2005) Stomatal patterning and differentiation by synergistic interactions of receptor kinases. Science **309**: 290-293

Silverman GA, Lomas DA (2007) *Molecular and Cellular Aspects of the Serpinopathies and Disorders in Serpin Activity*. Singapore, World Scientific [978-981-256-963-9]

Silverman GA, Whisstock JC, Bottomley SP, Huntington JA, Kaiserman D, Luke CJ, Pak SC, Reichhart J-M, Bird PI (2010) Serpins flex their muscle. I. Putting the clamps on proteolysis in diverse biological systems. J Biol Chem **285**: 24299-24305

Silverstein KAT, Moskal WA Jr, Wu HC, Underwood BA, Graham MA, Town CD, VandenBosch KA (2007) Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. Plant J **51**: 262-280

Stahl Y, Grabowski S, Bleckmann A, Kühnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hülsewede A, Felekyan S, Seidel CAM, Simon R (2013) Moderation of *Arabidopsis* root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes. Curr Biol **23**: 362-371

Stahl Y, Simon R (2012) Peptides and receptors controlling root development. Phil Trans R Soc B **367**: 1453-1460

Stahl Y, Wink RH, Ingram GC, Simon R (2009) A signaling module controlling the stem cell niche in *Arabidopsis* root meristems. Curr Biol **19**: 909-914

Stenvik G-E, Butenko MA, Urbanowicz BR, Rose JKC, Aalen RB (2006) Overexpression of *INFLORESCENCE DEFICIENT IN ABSCISSION* activates cell separation in vestigial abscission zones in *Arabidopsis*. Plant Cell **18**: 1467-1476

Stenvik G-E, Tandstad NM, Guo Y, Shi C-L, Kristiansen W, Holmgren A, Clark SE, Aalen RB, Butenko MA (2008) The EPIP peptide of INFLORESCENCE DEFICIENT IN ABSCISSION is sufficient to induce abscission in *Arabidopsis* through the receptor-like kinases HAESA and HAESA-LIKE2. Plant Cell **20**: 1805-1817

Stührwohldt N, Dahlke RI, Steffens B, Johnson A, Sauter M (2011) Phytosulfokine-α controls hypocotyl length and cell expansion in *Arabidopsis thaliana* through phytosulfokine receptor 1. PLoS ONE **6**: e21054

Suer S, Agusti J, Sanchez P, Schwarz M, Greb T (2011) *WOX4* imparts auxin responsiveness to cambium cells in *Arabidopsis*. Plant Cell **23**: 3247-3259

Sugano SS, Shimada T, Imai Y, Okawa K, Tamai A, Mori M, Hara-Nishimura I (2010) Stomagen positively regulates stomatal density in *Arabidopsis*. Nature **463**: 241-244

Takayama S, Shiba H, Iwano M, Shimosato H, Che F-S, Kai N, Watanabe M, Suzuki G, Hinata K, Isogai A (2000) The pollen determinant of self-incompatibility in *Brassica campestris*. Proc Natl Acad Sci USA **97**: 1920-1925

Uchida N, Lee JS, Horst RJ, Lai H-H, Kajita R, Kakimoto T, Tasaka M, Torii KU (2012) Regulation of inflorescence architecture by intertissue layer ligand--receptor communication between endodermis and phloem. Proc Natl Acad Sci USA **109**: 6337-6342

Uchida N, Tasaka M (2013) Regulation of plant vascular stem cells by endodermis-derived EPFL-family peptide hormones and phloem-expressed ERECTA-family receptor kinases. J Exp Bot **64**: 5335-5343

van de Sande K, Pawlowski K, Czaja I, Wieneke U, Schell J, Schmidt J, Walden R, Matvienko M, Wellink J, van Kammen A, Franssen H, Bisseling T (1996) Modification of phytohormone response by a peptide encoded by *ENOD40* of legumes and a nonlegume. Science **273**: 370-373

van der Hoorn RAL (2008) Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol **59**: 191-223

Van Norman JM, Breakfield NW, Benfey PN (2011) Intercellular communication during plant development. Plant Cell **23**: 855-864

Van Norman JM, Liberman LM, Benfey PN (2013a) *Arabidopsis* root: a model organ in plant genomics. In *Plant Roots: The Hidden Half*, 4th ed, A Eshel, T Beeckman (Eds). Boca Raton, FL, CRC Press, pp 2.1-2.17

Van Norman JM, Xuan W, Beeckman T, Benfey PN (2013b) To branch or not to branch: the role of pre-patterning in lateral root formation. Development **140**: 4301-4310

Vanneste S, Friml J (2009) Auxin: A trigger for change in plant development. Cell **136**: 1005-1016

Vanoosthuyse V, Miege C, Dumas C, Cock JM (2001) Two large *Arabidopsis thaliana* gene families are homologous to the *Brassica* gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response. Plant Mol Biol **46**: 17-34

Vanstraelen M, Benková E (2012) Hormonal interactions in the regulation of plant development. Annu Rev Cell Dev Biol **28**: 463-487

Vercammen D, Belenghi B, van de Cotte B, Beunens T, Gavigan J-A, De Rycke R, Brackenier A, Inzé D, Harris JL, Van Breusegem F (2006) Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. J Mol Biol **364**: 625-636

Wang G, Fiers M (2010) CLE peptide signaling during plant development. Protoplasma **240**: 33-43

Whitford R, Fernandez A, De Groodt R, Ortega E, Hilson P (2008) Plant CLE peptides from two distinct functional classes synergistically induce division of vascular cells. Proc Natl Acad Sci USA **105**: 18625-18630

Whitford R, Fernandez A, Tejos R, Cuéllar Pérez A, Kleine-Vehn J, Vanneste S, Drozdzecki A, Leitner J, Abas L, Aerts M, Hoogewijs K, Baster P, De Groodt R, Lin Y-C, Storme V, Van de Peer Y, Beeckman T, Madder A, Devreese B, Luschnig C, Hilson P (2012) GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses. Dev Cell **22**: 678-685

Wu J, Kurten EL, Monshausen G, Hummel GM, Gilroy S, Baldwin IT (2007) NaRALF, a peptide signal essential for the regulation of root hair tip apoplastic pH in *Nicotiana attenuata*, is required for root hair development and plant growth in native soils. Plant J **52**: 877-890

Yadav RK, Perales M, Gruel J, Girke T, Jönsson H, Reddy GV (2011) WUSCHEL protein movement mediates stem cell homeostasis in the *Arabidopsis* shoot apex. Genes Dev **25**: 2025-2030

Yamada M, Sawa S (2013) The roles of peptide hormones during plant root development. Curr Opin Plant Biol **16**: 56-61

Yamaguchi Y, Pearce G, Ryan CA (2006) The cell surface leucine-rich repeat receptor for *At*Pep1, an endogenous peptide elicitor in *Arabidopsis*, is functional in transgenic tobacco cells. Proc Natl Acad Sci USA **103**: 10104-10109

Yang H, Matsubayashi Y, Hanai H, Sakagami Y (2000) Phytosulfokine-α, a peptide growth factor found in higher plants: its structure, functions, precursor and receptors. Plant Cell Physiol **41**: 825-830

Yang S-L, Xie L-F, Mao H-Z, Puah C-S, Yang W-C, Jiang L, Sundaresan V, Ye D (2003) *TAPETUM DETERMINANT1* is required for cell specialization in the Arabidopsis anther. Plant Cell **15**: 2792-2804

Ye S, Goldsmith EJ (2001) Serpins and other covalent protease inhibitors. Curr Opin Struct Biol **11**: 740-745

Zhou A, Carrell RW, Huntington JA (2001) The serpin inhibitory mechanism is critically dependent on the length of the reactive center loop. J Biol Chem **276**: 27541-27547

Chapter 2

Proteases: Scissors processing signaling peptides in plants

This chapter is in preparation for submission.

Author’s contribution: The concept and structure of the manuscript was devised by S.G.; S.G. and P.H. wrote the manuscript with the help of A.F., F.V.B., and T.B.

Sarieh Ghorbani,[1,2] Ana Fernandez,[1,2] Tom Beeckman,[1,2] Frank Van Breusegem,[1,2] and Pierre Hilson[1,2,3,4*]

[1]Department of Plant Systems Biology, VIB, B-9052 Ghent, Belgium

[2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium

[3]INRA, UMR1318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

[4]AgroParisTech, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

[*]Corresponding author

Running head: proteases, processing signaling peptides

Proteases: scissors maturing signaling peptides

Sarieh Ghorbani, Ana Fernandez, Tom Beeckman, Frank Van Breusegem, Pierre Hilson

Introduction

In plants and animals alike, signaling peptides (SPs) participate in multicellular development through cell-to-cell communication. However, despite the increasing awareness of the importance of SPs in plant growth and development, little is known about their maturation process. This review summarizes the recent reports describing specific proteases required for the production of SP signals and discusses how these enzymes may modulate the downstream signaling pathways.

Signaling peptides have been associated with diverse developmental processes in plants such as apical meristem maintenance (CLAVATA3 (CLV3), GOLVEN/ROOT GROWTH FACTOR/CLE-LIKE (GLV/RGF/CLEL)), defense (systemins), floral organ abscission (INFLORESCENCE DEFICIENT IN ABSCISSION (IDA)), and stomatal patterning (EPIDERMAL PATTERNING FACTOR (EPF)) among others (Matsubayashi, 2011a, 2011b; Murphy et al., 2012; Aalen, 2013)].

Plant SPs belong to two large classes defined as cysteine-rich and cysteine-poor (Matsubayashi, 2011a). The mature signal molecule is released from the protein precursor by one or more proteolytic processing steps. Most SPs carry an amino- (N) terminal signal peptide cleaved upon secretion and necessary for their release into the endoplasmic reticulum (ER). Additional proteolytic cleavages may take place to yield the mature peptide encoded in a conserved motif defining different cysteine-poor SP subclasses (Hook et al., 2008; Matsubayashi, 2011a; Murphy et al., 2012). Furthermore, the C-terminal region of most cysteine-poor SPs is subjected to one or more post-translational modifications, namely, tyrosine sulfation (Amano et al., 2007; Matsuzaki et al., 2010), proline hydroxylation (Pearce et al., 2001b; Ito et al., 2006; Kondo et al., 2006; Ohyama et al., 2008; Ohyama et al., 2009; Matsuzaki et al., 2010) and hydroxyproline arabinosylation (Amano et al., 2007; Ohyama et al., 2009). The peptide

conformation is modulated by such modifications via steric and charge interactions that can alter its receptor binding affinity (Kondo et al., 2006; Ogawa, 2008; Ohyama et al., 2009; Matsubayashi, 2011b). In some cases, it has been shown that post-translationally modified peptides have higher biological activity than the unmodified version (Matsuzaki et al., 2010; Ohki et al., 2011; Shinohara and Matsubayashi, 2013).

Proteases may be in charge of general protein turnover or instead responsible for selective protein degradation (Davie and Neurath, 1955). They are called amino-, carboxy- and endopeptidases whether they hydrolyze peptide bonds at N-terminus, C-terminus, or within the body of a protein, respectively (Turk, 2006). The MEROPS database (release 9.9) lists 804 peptidases encoded in the *Arabidopsis thaliana* genome (Rawlings et al., 2012) and subdivided into families based on their catalytic mechanism. For instance, serine and cysteine proteases use serine and cysteine residues in their active site as nucleophile (Van der Hoorn, 2008). The proteases known to be involved in SP maturation are described in the next sections.

Signal peptide processing

The secretory signal peptide at the N-terminus of secreted proteins can be reliably predicted with the SignalP algorithm (Petersen et al., 2011). This signal-sorting sequence is present in almost all small SPs (Matsubayashi, 2011b; Murphy et al., 2012) (Fig 1A). It directs the SP through the secretory pathway via translocation of the precursor through the ER (Douglass et al., 1984). Once the translation of the pre-propeptides is initiated, the N-terminal signal peptide will be removed by a signal peptide peptidase (SPP) which is an aspartyl protease localized in the ER, yielding the propeptides. In *Arabidopsis* 6 genes encode for SPPs. They are collectively expressed in shoot and root meristem suggesting that they are involved in growth and organogenesis (Tamura et al., 2008; Han et al., 2009; Hoshi et al., 2013). Recently Hoshi et al., demonstrated *in vitro* cleavage of preprolactin, a human SPP substrate by AtSPP. Although preprolactin is not a native substrate for AtSPP, proteolytic process occurs in a similar manner to by which human SPP cleaves prolactin (Hoshi et al., 2013). Experimental data suggested an important role of AtSPP in male gametophyte development and pollen maturation in Arabidopsis. Although, the pollen in *spp* mutant plants is still viable but the grain fail to proceed to a stage that

allows proper germination during fertilization (Han et al., 2009). However, the native substrates targeted by SPP in plants have not been identified yet.

Biochemical evidence for SP–protease interactions

The simple comparison of SP native sequences with the encoding precursors indirectly demonstrates processing activity. Oligopeptides corresponding to the CEP1 (Ohyama et al., 2008), AtPEP1 (Huffaker et al., 2006), and PSY1 proproteins (Amano et al., 2007) have been described, while no protease associated with their processing has yet been pinpointed. On the other hand, specific proteases are known to participate in specific developmental programs while their native substrates remain unidentified (Neuteboom et al., 1999; Batchelor et al., 2000; Berger and Altmann, 2000; Zhao et al., 2000; Von Groll et al., 2002; Rautengarten et al., 2008). They may or may not be SP proproteins.

A closer connection between the molecular scissors-the proteases-and the signaling peptides, has been established with the biochemical analysis of plant extracts. The *Medicago truncatula* MtCLE36 propeptide and its soybean (*Glycine max*) orthologue, GmCLE34 share highly conserved sequences surrounding the typical C-terminal CLE domain, thus suggesting the presence of protease target sites within that region. Indeed, synthetic MtCLE36- and GmCLE34-derived oligopeptides encompassing the CLE motif together with N- and C-terminal extensions are proteolytically cleaved to yield the same 15 aa sequence when incubated with soybean xylem sap or *M. truncatula* extracellular fluids (Djordjevic et al., 2007; Kusumawati et al., 2008). Similar experiments point to the proteolytic release of the conserved CLE domain from the Arabidopsis CLV3 and CLE1 precursors following incubation with extracts from cauliflower (*Brassica oleracea*) and BY-2 tabacco cell suspension cultures (Ni and Clark, 2006). The corresponding bioactive mature peptide was also released from the IDA precursor after incubation with cauliflower extracts (Stenvik et al., 2008). A serine protease inhibitor (phenylmethanesulfonylfluoride, PMSF) abolishes the endoproteolytic cleavage of the CLE propeptide baits in both extracellular extracts, hinting at the involvement of subtilase activity (Ni et al., 2011). The mass spectrometry profiling of SP precursors following incubation in plant extracts also reveals series of byproducts resulting from the sequential removal of C-terminal

residues, a pattern compatible with the exoproteolytic activity of carboxypeptidases (Djordjevic et al., 2011).

Subtilisin-like serine proteases involved in SP maturation

The subtilase enzymes are present in archaea, bacteria and eukarya. Subtilases have a characteristic catalytic triad consisting of an aspartate, a histidine and a serine residue, which together generate the nucleophilic potential that is necessary for substrate binding ability. (Dodson and Wlodawer, 1998). In Arabidopsis, they form a large family of 56 members subdivided into six distinct subfamilies based on homology between the deduced full-length amino acid sequences (Rautengarten et al., 2005). Converging evidence shows that subtilisin-like serine proteases, hereafter referred to as subtilases, are involved in the processing of small plant signaling propeptides leading to the release of the bioactive signal encoded in their C-terminal region (Seidah and Chretien, 1999; Rholam and Fahy, 2009; Schaller et al., 2012).

a. Phytosulfokines (PSK)

Phytosulfokines (PSKs) were originally isolated from conditioned medium of asparagus (*Asparagus officinallis*) mesophyll cells (Matsubayashi and Sakagami, 1996). Genes coding for PSK precursors have been identified in all higher plants, including five in Arabidopsis (Yang et al., 2000; Yang et al., 2001; Matsubayashi and Sakagami, 2006). PSKs are considered ubiquitous growth promoting factors associated with both cell elongation and cell proliferation (Matsubayashi and Sakagami, 1996; Kutschmar et al., 2009; Stuhrwohldt et al., 2011; Hartmann et al., 2013). However, in *Zinnia elegans*, PSK can also stimulate the differentiation of dispersed mesophyll cells into tracheary elements (Matsubayashi et al., 1999). PSK precursors are approximately 80 amino-acids in length and contain a signal peptide in their N-terminal region (Matsubayashi, 2011a). Mature PSKs are oligopeptides of five amino acids (aa) (YIYTQ) with both tyrosine residues sulfated (Matsubayashi and Sakagami, 1996).

It has been shown that a subtilisin-like serine proteases, *AtSBT1.1* is required for efficient shoot regeneration (Lall et al., 2004). As PSKs stimulate callus formation in tissue culture and carry dibasic residues typical of subtilase target sites in their precursor sequences, Sirvastava et al. studied the possible role of AtSBT1.1 in the processing of these SPs (Srivastava et al., 2008).

An epitope-tagged version of the PSK4 precursor was processed in transgenic plants under callus inducing condition, but no proteolytic activity was observed in non-callus inducing condition. Furthermore, the epitope-tagged PSK4 was not proteolytically cleaved when expressed into the *sbt1.1* knockout mutant. The biochemical activity of SBT1.1 was confirmed *in vitro* with synthetic PSK4 fluorogenic peptide substrate. The mass spectrometry analysis of proteolytic byproducts revealed that the enzyme cleaves the propeptide three aa upstream of the mature peptide sequence and few aa downstream of its di-basic motif (RRSLVL↓HTDY(NO2)D-OH). Substitution of residues with alanine at the protease recognition site hinders the proteolytic process. And while AtSBT1.1 could process the other members of the family, it shows preference for PSK4 (Srivastava et al., 2008).

Basic residues may define the initial processing sites but do not always directly mark the border of the mature peptide (Matsubayashi, 2011b). For example, the cleavage of the PSK4 precursor occurs three aa upstream of the mature peptide sequence (Srivastava et al., 2008). Furthermore, all PSK precursor sequences hold a few extra aa downstream of the mature peptideS (Yang et al., 2001). Therefore, additional processing steps are most likely required to produce the bioactive signal, either through endoproteolysis or exoproteolytic trimming (Matsubayashi and Sakagami, 2006; Ni et al., 2011).

b. Rapid alkalinization factors (RALF)

Rapid alkalinization factor (RALF) was first isolated from tobacco leaves as a molecule that triggers fast and strong pH hikes in tobacco suspension-cultured cells, that stimulates a mitogen-activated protein kinase (MAPK), and that arrests root growth (Pearce et al., 2001b) The cysteine-rich RALF-coding genes are found throughout the plant kingdom and Arabidopsis counts 34 *RALF* homologues (Olsen et al., 2002). The *RALF* and *RALF-like (RAFL)* genes are expressed in most tissues including roots, shoots, leaves and flowers (Pearce et al., 2001b; Germain et al., 2005; Combier et al., 2008; Bedinger et al., 2010), and are suspected to regulate diverse developmental modules, such as root development and root hair growth (Pearce et al., 2001b; Wu et al., 2007), pollen elongation (Covey et al., 2010), and possibly nodulation in *Medicago trunculata* [*MtRALF1;* (Combier et al., 2008)]. Similar to PSKs, RALFs possess a N-

terminal signal peptide indicating they are secreted (Pearce et al., 2001b) and a conserved dibasic site that could be targeted for subtilase proteolysis to generate the mature peptide.

AtRALF23 is downregulated by brassinosteroids (BR) and is thought to be a negative regulator of BR-mediated growth-promoting effects (Nemhauser et al., 2004; Srivastava et al., 2008). In the *AtRALF23* 138-aa precursor, a dibasic site is located immediately upstream of the predicted mature peptide. Prompted by this observation, Srivastava et al. showed that the Arabidopsis *AtRALF23* overproduction phenotypes – dwarf bushy plants – is suppressed in the *sbt6.1* null mutant line and that no processed AtRALF23 peptide was detected in these plants (Srivastava et al., 2009). Again, the substitution of two arginines with a glycine residue in the protease recognition site (RRIL) prevented *in planta* cleavage and the plant-purified AtSBT6.1 enzyme could process the AtRALF23 *in vitro*, thereby confirming that the propeptide is a direct target of the subtilase (Srivastava et al., 2009). Unlike AtSBT1.1 action on the AtPSK4 propeptide (Srivastava et al., 2008), the peptide bond cleaved by AtSBT6.1 in AtRALF23 is positioned exactly at the N-terminus of the mature peptide sequence. AtSBT6.1 has been localized in the Golgi apparatus (Liu et al., 2007) and may therefore process AtRALF23 in that compartment on its way to the apoplast through the secretory path (Srivastava et al., 2009).

c. GOLVEN/Root meristem growth factor/CLE like peptides (GLV/RGF/CLEL)

The GLV/RGF/CLE-like SP family were co-discovered by three groups through independent *in silico* studies (Matsuzaki et al., 2010; Meng et al., 2012; Whitford et al., 2012). For clarity, these peptides are referred to hereafter according to their GLV nomenclature. The family is found in all higher plants examined so far and consists of 11 members in Arabidopsis. GLV peptides have been implicated in root apical meristem maintenance, root and hypocotyl gravitropic responses, lateral root formation and root hair development (Fernandez et al., 2013a; Fernandez et al., 2013b). Similar to PSKs, GLVs are encoded as larger pre-propeptides, subjected to multiple proteolytic cleavages as well as posttranslational modifications to yield mature bioactive molecules. The sequence of the native mature peptides has been defined by mass spectrometry for four members, including GLV1, a 14-aa oligopeptide that carries a sulfated tyrosine and one or two hydroxylated prolines (Matsuzaki et al., 2010; Whitford et al., 2012).

Driven by results demonstrating subtilase-processing of the plant PSKs and RALFs, Ghorbani et al. screened for possible suppression of *GLV1* gain-of-function phenotype – agravitropic roots – in Arabidopsis subtilases T-DNA knockout lines. They tested mutants for 55 of the 56 identified subtilase genes (Rautengarten et al., 2005) by transformation with a *GLV1* overexpression transgene. The suppressor screen revealed that the *SBT6.1* and *SBT6.2* subtilase genes are essential for the maturation and activation of the GLV1 peptide (Ghorbani et al., in preparation). While the SBT6.1 and SBT6.2 proteins are more closely related to each other than to any other Arabidopsis subtilases, they are not genetically redundant: recessive null mutations in either gene resulted in undistinguishable phenotypes and the *sbt6.1-1 sbt2* double mutant did not exhibit an additive phenotype, suggesting that are both involved in GLV1 processing but at different steps.

In vitro protease assays confirmed that the plant-purified SBT6.1 enzyme cleaves the GLV1 precursor peptides at sites resembling the canonical subtilases recognition sequences, RXXL and RXLX (Rholam and Fahy, 2009; Schaller et al., 2012). Nevertheless, these sites are located several residues upstream of the native peptide, indicating that SBT6.1 is not sufficient to produce the mature peptide detected in plant tissues (Whitford et al., 2012). Besides, three extra aa encoded in the propeptide are eventually removed from its C-terminal end to result in the mature GLV1 peptide. Mammalian counterpart of SBT6.2, TPP II is known as an exopeptidase, which removes tripeptides from the amino-termini of larger peptides (Barr, 1991b). The additional processing steps are probably catalyzed by other proteases, possibly involving SBT6.2.

Since 10 out of 11 GLV precursors carry at least one sequence compatible with subtilase processing, SBT6.1 may cleave multiple members of the GLV family (Ghorbani et al., in preparation). In fact, this is likely the case considering that *amiRglv1* silenced lines do not show any discernible developmental mutant phenotypes in contrast to the *sbt6.1-1* and *sbt6.2* mutants. However, the short hypocotyls of the *sbt6* mutants may be caused by defective processing of GLVs, but also other proteins, including SPs such as RALFs.

The analysis of SBT6.1 function revealed an unexpected additional level of peptide processing regulation. Tandem-affinity purification experiments confirmed by *in vivo* bimolecular fluorescence complementation assays showed that SBT6.1 forms a protein complex with the Serpin1 protease inhibitor. Serpins are suicide inhibitors that form an irreversible

covalent bond with their targets (Huntington et al., 2000). While serpins are the largest and most broadly distributed superfamily of protease inhibitors, their function in plants remains poorly understood (Huntington et al., 2000; Law et al., 2006). The *sbt6* short-hypocotyl loss-of-funtion phenotype is phenocopied by *Serpin1* overexpression, and the Serpin1 protein purified from bacterial extracts prevents the proteolytic cleavage of GLV1 precursor fragments by SBT6.1 *in vitro*. Altogether, these results suggest that GLV1, SBT6.1 and Serpin1 form a signaling module controlling hypocotyl development in Arabidopsis (Ghorbani et al., in preparation).

Exopeptidases involved in SP maturation

SP gene overexpression phenotypes can also be recapitulated by treatments of plant tissues with the corresponding synthetic oligopeptides (Fiers et al., 2005; Whitford et al., 2008). It was therefore straightforward to notice that the addition or removal of a single N- or C-terminal residue to the experimentally verified mature peptide may result in an inactive peptide (Kondo et al., 2006; Djordjevic et al., 2011), thereby highlighting the importance of proper peptide processing at both ends of the conserved SP conserved motif. However, as already pointed out, the known subtilase endoproteolytic byproducts of AtPSK4 and GLV1 do not match the native SP molecules extracted from plant tissues or conditioned media. Thus, additional proteases are part of the maturation pipelines releasing the bioactive SPs.

One of them is *SOL1* whose mutation suppresses *CLE19* overexpression phenotypes (Casamitjana-Martinez et al., 2003). SOL1 is a putative transmembrane Zn^{2+}-carboxypeptidase partly homologous to animal carboxypeptidases D and E cleaving off C-terminal arginine and lysine residues from neuropeptides and prohormones (Fricker, 1988; Greene et al., 1992; Sidyelyeva and Fricker, 2002). Recently, Tamaki et al. showed that SOL1 is necessary for trimming a C-terminal arginine off the CLE19 SP in Arabidopsis, most likely in endosomal compartments, to yield the active CLE19 peptide. In biochemical assays, SOL1 has lysine as well as arginine exopeptidase activity (Tamaki et al., 2013)(Fig. 1B). Since it is expressed ubiquitously in different tissues, SOL1 may be involved in processing other SPs. For example, CLE family members (CLE14, CLE20, CLE22, CLE42) as well as GLV2 (Olsen and Skriver, 2003; Whitford et al., 2012) carry an arginine or lysine residue, immediately after the mature SP sequence in the precursor protein (Tamaki et al., 2013).

Moreover, mammalian counterpart of SBT6.2, TPP II is known as an exopeptidase, which removes tripeptides from the amino-termini of larger peptides (Barr, 1991b). Possibly SBT6.2 plays the same biological function, but that requires further investigation.

Signaling peptide processing: housekeeping or refined programming?

The discovery of multiple proteases with specific SP targets has eroded the view that these processing enzymes are mainly taking care of general protein turnover (Van der Hoorn, 2008; Schaller et al., 2012). In fact, the initial phenotypic survey of null Arabidopsis subtilase mutants was somewhat disappointing because only a few displayed readily recognizable phenotypes (Rautengarten et al., 2005). This can be interpreted as a sign of genetic redundancy or instead that they act very specifically, for example upstream of SP signaling pathways. Their connection with SP processing has shed new light on their functions and drew our attention on particular developmental processes controlling cell fate, cell expansion and cell division.

Proteolytic processing may modulate SP action in several ways. First, pre-propeptides can be stored as a stable and inactive pool rapidly and massively triggered by cleavage when or where it meets its cognate protease. Second, complex gradients of morphogenic signals may result from the overlap between fields of diffusing propeptide and protease, each released in the extracellular space from defined cell types or at specific times (Wheeler and Irving, 2010; Wheeler and Irving, 2012). Third, as most of the signaling peptides are secreted, the production of the signaling peptide initially as a pre-propeptide may provide the necessary features of the secretion within the structure of precursors such as holding signal peptide. Finally, proteolytic cleavage may modulate the speed at which SPs move through the apoplast, the smaller the peptide size the higher the diffusion rate (Wheeler and Irving, 2010; Srivastava and Howell, 2012; Shinohara and Matsubayashi, 2013).

To our surprise, the Arabidopsis SBT6.1 subtilase has been found to process two SPs from different families: RALF23 that elicits an alkalinisation cell response and reduces hypocotyl and root growth (Srivastava et al., 2009); GLV1 that modulates auxin fluxes and, in contrast, promotes cell elongation (Ghorbani et al., in preparation). Thus, the same processing enzyme appears to control the production of peptide signals with opposite effects on organ growth. We

speculate that SBT6.1 and other proteases may thereby act as integrators of competing modules (Fig. 1B).

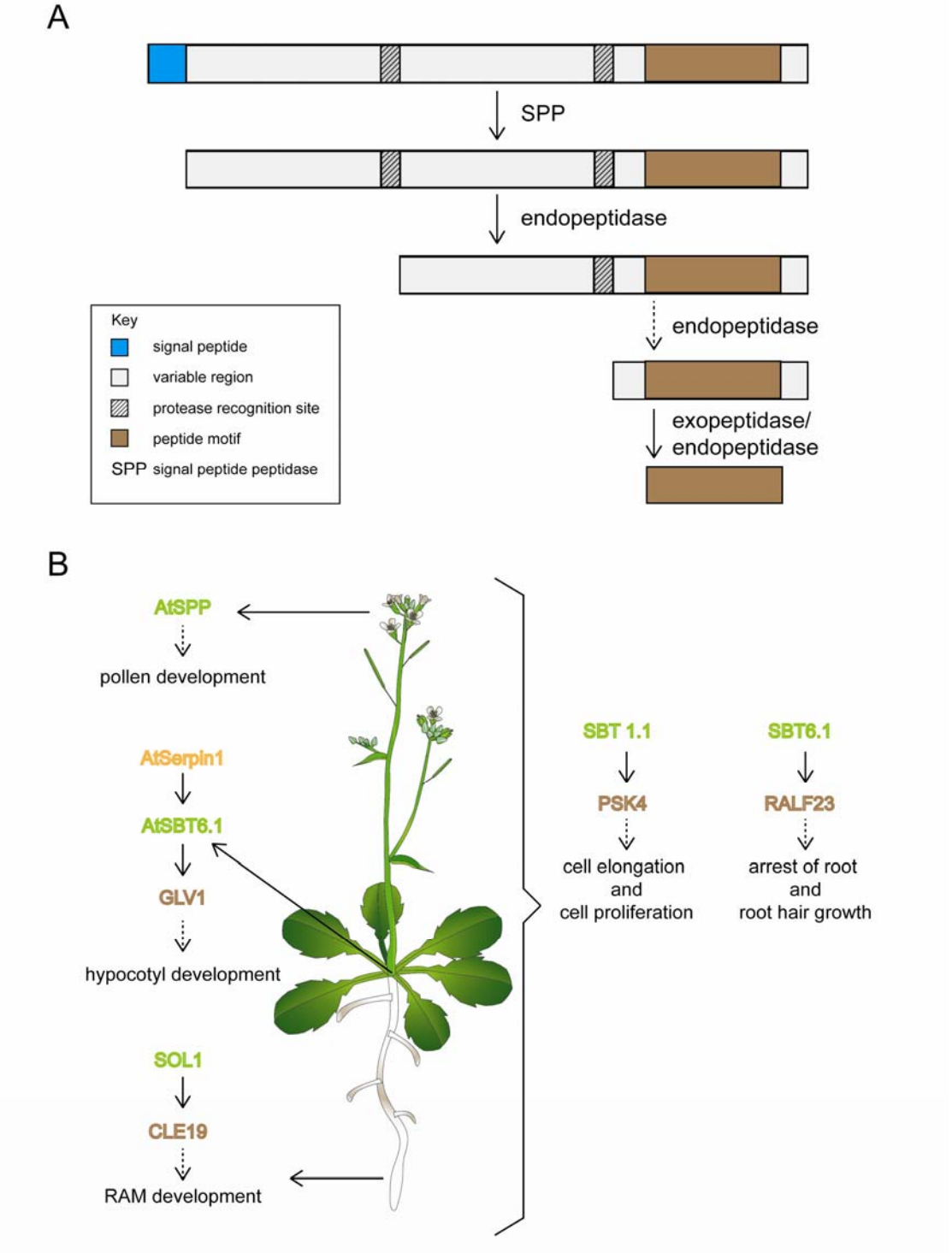


Figure 1. Processing of Small Signaling Peptides.

A., Following entry of the full-length prepropeptide to the secretory pathway, the N-terminal signal peptide is cleaved by a signal peptidase. Later on one or more proteolytic processes are necessary to yield the exact mature peptide sequence or with some extra flanking amino acids. Additional endoproteolytic process or exoproteolytic trimming is needed to produce mature peptide. **B.**, Schematic representation of the protease and signaling peptide-depended developmental features in *Arabidopsis* reported so far. Proteases are indicated in green text, signaling peptides in brown and inhibitors in orange.

Future perspectives

Because the proteolytic process is irreversible, protease activity must be strictly monitored (Van der Hoorn, 2008; Fluhr et al., 2012). Indeed, protease and peptide genes show remarkable spatiotemporal expression patterns and seem to act in selected subcellular compartments. However, the overlap between their expression patterns or subcellular compartment location remains to be fully characterized. Overall, our understanding of protease-SP interactions remains very limited considering the large repertoire of proteases [~800 in Arabidopsis; (Rawlings et al., 2012)] and small secreted peptides [>1,000; (Lease and Walker, 2006; Ohyama et al., 2008)] encoded in plant genomes.

So far, protease-SP interactions have essentially been investigated through (i) forward genetic screens based on the suppression of SP gain-of-function phenotypes and (ii) the profiling of proteolytic cleavage of SP precursors, either in plant tissues or *in vitro* with plant extracts or purified enzymes. In parallel, the search for protease targets, possibly including SPs, can now be performed at a large-scale either through comparative proteomics/peptidomics mass spectrometry (Farrokhi et al., 2008; Tsiatsiani et al., 2012) or with microarray-based proteolytic profiling assay incorporating fluorogenic peptides (Gosalia et al., 2005a; Gosalia et al., 2005b). Such proteomics approaches combined with bioinformatics tools greatly facilitate the analysis of the protein degradomes, leading to the identification of the protease optimal target sites (Turk and Cantley, 2003). They are also useful to identify specific inhibitors of proteolytic processing activities and, thereby, to investigate the role of proteases in controlling development (Cantley and Turk, 2003). Furthermore, yeast two-hybrid, phage display and immobilized protein arrays can also reveal biologically relevant substrates for proteases (Deperthes, 2002; Schilling and Overall, 2007). In the emerging research field focusing on plant peptide signaling, proprotein processing has been recognized as a crucial step for the production of the bioactive signal. The ongoing characterization of peptide-protease interactions is necessary for a comprehensive understanding of cell-to-cell signaling in plant development.

References

Aalen RB (2013) Maturing peptide open for communications. J Exp Bot **64**: 5231-5235

Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. Proc Natl Acad Sci USA **104**: 18333-18338

Barr PJ (1991b) Mammalian subtilisins: the long-sought dibasic processing endoproteases. Cell **66**: 1-3

Batchelor AK, Boutilier K, Miller SS, Labbé H, Bowman L, Hu M, Johnson DA, Gijzen M, Miki BLA (2000) The seed coat-specific expression of a subtilisin-like gene, *SCS1*, from soybean. Planta **211**: 484-492

Bedinger PA, Pearce G, Covey PA (2010) RALFs: peptide regulators of plant growth. Plant Signal Behav **5**: 1342-1346

Berger D, Altmann T (2000) A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. Genes Dev **14**: 1119-1131

Casamitjana-Martínez E, Hofhuis HF, Xu J, Liu C-M, Heidstra R, Scheres B (2003) Root-specific *CLE19* overexpression and the *sol1/2* suppressors implicate a CLV-like pathway in the control of *Arabidopsis* root meristem maintenance. Curr Biol **13**: 1435-1441

Combier J-P, Küster H, Journet E-P, Hohnjec N, Gamas P, Niebel A (2008) Evidence for the involvement in nodulation of the two small putative regulatory peptide-encoding genes *MtRALFL1* and *MtDVL1*. Mol Plant-Microbe Interact **21**: 1118-1127

Covey PA, Subbaiah CC, Parsons RL, Pearce G, Lay FT, Anderson MA, Ryan CA, Bedinger PA (2010) A pollen-specific RALF from tomato that regulates pollen tube elongation. Plant Physiol **153**: 703-715

Davie EW, Neurath H (1955) Identification of a peptide released during autocatalytic activation of trypsinogen. J Biol Chem **212**: 515-529

Deperthes D (2002) Phage display substrate: a blind method for determining protease specificity. Biol Chem **383**: 1107-1112

Djordjevic MA, Oakes M, Li DX, Hwang CH, Hocart CH, Gresshoff PM (2007) The *Glycine max* xylem sap and apoplast proteome. J Proteome Res **6**: 3771-3779

Djordjevic MA, Oakes M, Wong CE, Singh M, Bhalla P, Kusumawati L, Imin N (2011) Border sequences of *Medicago truncatula* CLE36 are specifically cleaved by endoproteases common to the extracellular fluids of *Medicago* and soybean. J Exp Bot **62**: 4649-4659

Dodson G, Wlodawer A (1998) Catalytic triads and their relatives. Trends Biochem Sci **23**: 347-352

Douglass J, Civelli O, Herbert E (1984) Polyprotein gene expression: generation of diversity of neuroendocrine peptides. Annu Rev Biochem **53**: 665-715

Farrokhi N, Whitelegge JP, Brusslan JA (2008) Plant peptides and peptidomics. Plant Biotechnol J **6**: 105-134

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, Hilson P (2013) Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol **161**: 954-970

Fernandez A, Hilson P, Beeckman T (2013b) GOLVEN peptides as important regulatory signalling molecules of plant development. J Exp Bot **64**: 5263-5268

Fluhr R, Lampl N, Roberts TH (2012) Serpin protease inhibitors in plant biology. Physiol Plant **145**: 95-102

Fricker LD (1988) Carboxypeptidase E. Annu Rev Physiol **50**: 309-321

Germain H, Chevalier E, Caron S, Matton DP (2005) Characterization of five RALF-like genes from *Solanum chacoense* provides support for a developmental role in plants. Planta **220**: 447-454

Gosalia DN, Salisbury CM, Ellman JA, Diamond SL (2005a) High throughput substrate specificity profiling of serine and cysteine proteases using solution-phase fluorogenic peptide microarrays. Mol Cell Proteomics **4**: 626-636

Gosalia DN, Salisbury CM, Maly DJ, Ellman JA, Diamond SL (2005b) Profiling serine protease substrate specificity with solution phase fluorogenic peptide microarrays. Proteomics **5**: 1292-1298

Greene D, Das B, Fricker LD (1992) Regulation of carboxypeptidase E. Effect of pH, temperature and Co^{2+} on kinetic parameters of substrate hydrolysis. Biochem J **285**: 613-618

Han S, Green L, Schnell DJ (2009) The signal peptide peptidase is required for pollen function in Arabidopsis. Plant Physiol **149**: 1289-1301

Hartmann J, Stührwohldt N, Dahlke RI, Sauter M (2013) Phytosulfokine control of growth occurs in the epidermis, is likely to be non-cell autonomous and is dependent on brassinosteroids. Plant J **73**: 579-590

Hook V, Funkelstein L, Lu D, Bark S, Wegrzyn J, Hwang S-R (2008) Proteases for processing proneuropeptides into peptide neurotransmitters and hormones. Annu Rev Pharmacol Toxicol **48**: 393-423

Hoshi M, Ohki Y, Ito K, Tomita T, Iwatsubo T, Ishimaru Y, Abe K, Asakura T (2013) Experimental detection of proteolytic activity in a signal peptide peptidase of *Arabidopsis thaliana*. BMC Biochemistry **14**: 16

Huffaker A, Pearce G, Ryan CA (2006) An endogenous peptide signal in *Arabidopsis* activates components of the innate immune response. Proc Natl Acad Sci USA **103**: 10098-10103

Huntington JA, Read RJ, Carrell RW (2000) Structure of a serpin--protease complex shows inhibition by deformation. Nature **407**: 923-926

Ito Y, Nakanomyo I, Motose H, Iwamoto K, Sawa S, Dohmae N, Fukuda H (2006) Dodeca-CLE peptides as suppressors of plant stem cell differentiation. Science, 313, 842-845

Kondo T, Sawa S, Kinoshita A, Mizuno S, Kakimoto T, Fukuda H, Sakagami Y (2006) A plant peptide encoded by *CLV3* identified by in situ MALDI-TOF MS analysis. Science **313**: 845-848

Kusumawati L, Imin N, Djordjevic MA (2008) Characterization of the secretome of suspension cultures of *Medicago* species reveals proteins important for defense and development. J Proteome Res **7**: 4508-4520

Kutschmar A, Rzewuski G, Stührwohldt N, Beemster GTS, Inzé D, Sauter M (2009) PDK-α promotes root growth in Arabidopsis. New Phytol **181**: 820-831

Kutschmar A, Rzewuski G, Stührwohldt N, Beemster GTS, Inzé D, Sauter M (2009) PSK-α promotes root growth in Arabidopsis. New Phytol **181**: 820-831

Lall S, Nettleton D, DeCook R, Che P, Howell SH (2004) Quantitative trait loci associated with adventitious shoot formation in tissue culture and the program of shoot development in Arabidopsis. Genetics **167**: 1883-1892

Law RHP, Zhang Q, McGowan S, Buckle AM, Silverman GA, Wong W, Rosado CJ, Langendorf CG, Pike RN, Bird PI, Whisstock JC (2006) An overview of the serpin superfamily. Genome Biol **7**: 216

Lease KA, Walker JC (2006) The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol **142**: 831-838

Liu J-X, Srivastava R, Che P, Howell SH (2007) Salt stress responses in Arabidopsis utilize a signal transduction pathway related to endoplasmic reticulum stress signaling. Plant J **51**: 897-909

Matayoshi ED, Wang GT, Krafft GA, Erickson J (1990) Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. Science **247**: 954-958

Matsubayashi Y (2011a) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol **52**: 5-13

Matsubayashi Y (2011b) Small post-translationally modified peptide signals in Arabidopsis. Arabidopsis Book 9, e0150.

Matsubayashi Y, Sakagami Y (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. Proc Natl Acad Sci USA **93**: 7623-7627

Matsubayashi Y, Sakagami Y (2006) Peptide hormones in plants. Annu Rev Plant Biol **57**: 649-674

Matsubayashi Y, Takagi L, Omura N, Morita A, Sakagami Y (1999) The endogenous sulfated pentapeptide phytosulfokine-α stimulates tracheary element differentiation of isolated mesophyll cells of zinnia. Plant Physiol **120**: 1043-1048

Matsuzaki Y, Ogawa-Ohnishi M, Mori A, Matsubayashi Y (2010) Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. Science **329**: 1065-1067

Meng L, Buchanan BB, Feldman LJ, Luan S (2012) CLE-like (CLEL) peptides control the pattern of root growth and lateral root development in *Arabidopsis*. Proc Natl Acad Sci USA **109**: 1760-1765

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in *Arabidopsis* development: how cells communicate over a short distance. Plant Cell **24**: 3198-3217

Nemhauser JL, Mockler TC, Chory J (2004) Interdependency of brassinosteroid and auxin signaling in *Arabidopsis*. PLoS Biol **2**: e258 (1460-1471)

Neuteboom LW, Veth-Tello LM, Clijdesdale OR, Hooykaas PJJ, van der Zaal BJ (1999) A novel subtilisin-like protease gene from *Arabidopsis thaliana* is expressed at sites of lateral root emergence. DNA Res **6**: 13-19

Ni J, Clark SE (2006) Evidence for functional conservation, sufficiency, and proteolytic processing of the CLAVATA3 CLE domain. Plant Physiol **140**: 726-733

Ni J, Guo Y, Jin H, Hartsell J, Clark SE (2011) Characterization of a CLE processing activity. Plant Mol Biol **75**: 67-75

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y (2008) *Arabidopsis* CLV3 peptide directly binds CLV1 ectodomain. Science **319**: 294-294 [Err. Science 319, 901]

Ohki S, Takeuchi M, Mori M (2011) The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. Nat Commun **2**: 512

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis. Plant J **55**: 152-160

Ohyama K, Shinohara H, Ogawa-Ohnishi M, Matsubayashi Y (2009) A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*. Nat Chem Biol **5**: 578-580

Olsen AN, Mundy J, Skriver K (2002) Peptomics, identification of novel cationic *Arabidopsis* peptides with conserved sequence motifs. In Silico biol **2**: 441-451

Olsen AN, Skriver K (2003) Ligand mimicry? Plant-parasitic nematode polypeptide with similarity to CLAVATA3. Trends Plant Sci **8**: 55-57

Pearce G, Moura DS, Stratmann J, Ryan CA Jr (2001b) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proc Natl Acad Sci USA **98**: 12843-12847

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8**: 785-786

Rautengarten C, Steinhauser D, Büssis D, Stintzi A, Schaller A, Kopka J, Altmann T (2005) Inferring hypotheses on functional relationships of genes: Analysis of the *Arabidopsis thaliana* subtilase gene family. PLoS Comput Biol **1**: e40

Rautengarten C, Usadel B, Neumetzler L, Hartmann J, Büssis D, Altmann T (2008) A subtilisin-like serine protease essential for mucilage release from Arabidopsis seed coats. Plant J **54**: 466-480

Rawlings ND, Barrett AJ, Bateman A (2012) *MEROPS*: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res **40**: D343-D350

Rholam M, Fahy C (2009) Processing of peptide and hormone precursors at the dibasic cleavage sites. Cell Mol Life Sci **66**: 2075-2091

Schaller A, Stintzi A, Graff L (2012) Subtilases - versatile tools for protein turnover, plant development, and interactions with the environment. Physiol Plant **145**: 52-66

Schilling O, Overall CM (2007) Proteomic discovery of protease substrates. Curr Opin Chem Biol **11**: 36-45

Seidah NG, Chrétien M (1999) Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. Brain Res **848**: 45-62

Shinohara H, Matsubayashi Y (2013) Chemical synthesis of Arabidopsis CLV3 glycopeptide reveals the impact of hydroxyproline arabinosylation on peptide conformation and activity. Plant Cell Physiol **54**: 369-374

Sidyelyeva G, Fricker LD (2002) Characterization of *Drosophila* carboxypeptidase D. J Biol Chem **277**: 49613-49620

Srivastava R, Howell SH (2012) Processing of peptides. In *Plant Signaling Peptides*, Signaling and Communication in Plants Series, Vol 16, HR Irving, C Gehring (Eds). Berlin, Springer, pp 199-215

Srivastava R, Liu J-X, Guo H, Yin Y, Howell SH (2009) Regulation and processing of a plant peptide hormone, AtRALF23, in Arabidopsis. Plant J **59**: 930-939

Srivastava R, Liu J-X, Howell SH (2008) Proteolytic processing of a precursor protein for a growth-promoting peptide by a subtilisin serine protease in Arabidopsis. Plant J **56**: 219-227

Stührwohldt N, Dahlke RI, Steffens B, Johnson A, Sauter M (2011) Phytosulfokine-α controls hypocotyl length and cell expansion in *Arabidopsis thaliana* through phytosulfokine receptor 1. PLoS ONE **6**: e21054

Tamaki T, Betsuyaku S, Fujiwara M, Fukao Y, Fukuda H, Sawa S (2013) SUPPRESSOR OF LLP1 1-mediated C-terminal processing is critical for CLE19 peptide activity. Plant J **76**: 970-81

Tamura T, Asakura T, Uemura T, Ueda T, Terauchi K, Misaka T, Abe K (2008) Signal peptide peptidase and its homologs in *Arabidopsis thaliana* -- plant tissue-specific expression and distinct subcellular localization. FEBS J **275**: 34-43

Tsiatsiani L, Gevaert K, Van Breusegem F (2012) Natural substrates of plant proteases: how can protease degradomics extend our knowledge? Physiol Plant **145**: 28-40

Turk B (2006) Targeting proteases: successes, failures and future prospects. Nat Rev Drug Discov **5**: 785-799

van der Hoorn RAL (2008) Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol **59**: 191-223

von Groll U, Berger D, Altmann T (2002) The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during Arabidopsis stomatal development. Plant Cell **14**: 1527-1539

Wheeler JI, Irving HR (2010) Evolutionary advantages of secreted peptide signalling molecules in plants. Funct Plant Biol **37**: 382-394

Wheeler JI, Irving HR (2012) Plant peptide signaling: an evolutionary adaptation. In *Plant Signaling Peptides*, Signaling and Communication in Plants Series, Vol 16, HR Irving, C Gehring (Eds). Berlin, Springer, pp 1-23

Whitford R, Fernandez A, Tejos R, Cuéllar Pérez A, Kleine-Vehn J, Vanneste S, Drozdzecki A, Leitner J, Abas L, Aerts M, Hoogewijs K, Baster P, De Groodt R, Lin Y-C, Storme V, Van de Peer Y, Beeckman T, Madder A, Devreese B, Luschnig C, Hilson P (2012) GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses. Dev Cell **22**: 678-685

Wu J, Kurten EL, Monshausen G, Hummel GM, Gilroy S, Baldwin IT (2007) NaRALF, a peptide signal essential for the regulation of root hair tip apoplastic pH in *Nicotiana attenuata*, is required for root hair development and plant growth in native soils. Plant J **52**: 877-890

Yang H, Matsubayashi Y, Hanai H, Sakagami Y (2000) Phytosulfokine-α, a peptide growth factor found in higher plants: its structure, functions, precursor and receptors. Plant Cell Physiol **41**: 825-830

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y (2001) Diversity of Arabidopsis genes encoding precursors for phytosulfokine, a peptide growth factor Plant Physiol **127**: 842-851

Zhao C, Johnson BJ, Kositsup B, Beers EP (2000) Exploiting secondary growth in Arabidopsis. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. Plant Physiol **123**: 1185-1196

Outlook

During the last decade, our knowledge about plant signaling peptides has progressed considerably and a number of signaling peptide families have been discovered and partially characterized. However, considering the thousands of small open reading frames in *Arabidopsis* that potentially code for small signaling peptides, a lot more remains to be unraveled. In this regard, we performed an *in silico* research in an attempt to look for new probable signaling peptides implicated in lateral root development.

Furthermore, signaling peptides often are part of a complex network, of which several elements collaborate precisely to relay the signal in signaling peptide pathways. Processing enzymes are one of these key features in the production of the mature peptides. Our work aimed at finding the potential enzyme(s) involved in the maturation of the GLV precursor(s) as well as additional processing regulators.

Chapter 3

A signaling peptide-processing module controls cell elongation.

This chapter is in preparation for submission.

Author contributions: S.G., A.F., F.V.B., and P.H., designed the experiments; S.G. performed the experiments related to figures 1, 2, 3, S1, S2; A.I. generated the Serpin1 lines for these experiments; K.H: performed the experiments related to figure 4, S3, T.P. and A.I. performed the experiments related to figure 5; D.E. performed TAP experiments; S.G., K.H., A.F., D.E., G.D.J., T.B., A.M., F.V.B., and P.H., analyzed the data; S.G. and P.H. wrote the manuscript

A signaling peptide-processing module controls cell elongation[1][W]

Sarieh Ghorbani,[a,b] Kurt Hoogewijs,[c] Tamara Pečenková,[a,b] Ana Fernandez,[a,b] Annelies Inzé,[a,b] Dominique Eeckhout,[a,b] Dorota Kawa,[a,b] Geert De Jaeger,[a,b] Tom Beeckman,[a,b] Annemieke Madder,[c] Frank Van Breusegem,[a,b] and Pierre Hilson[a,b,d,e,*]

[a]Department of Plant Systems Biology, VIB, B-9052 Ghent, Belgium (S.G., T.P., A.F., A.I., D.E., D.K., G.D.J., T.B., F.V.B., and P..H.); [b]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium (S.G., T.P., A.F., A.I., D.E., D.K., G.D.J., T.B., F.V.B., and P..H.); [c]Department of Organic Chemistry, Ghent University, B-9000 Ghent, Belgium (K.H., A.M.); [d]Institut National de la Recherche Agronomique, Unité Mixte de Recherche 318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France (P.H.); [e]AgroParisTech, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France (P.H.)

Footnotes

[1] This work was supported by the Sixth Framework Programme of the European Commission (Integrated Project AGRON-OMICS [LSHG-CT-2006-037704]), the Interuniversity Attraction Poles Programme (IAP P7/29), initiated by the Belgian Science Policy Office, and the Research Foundation-Flanders (predoctoral fellowship to A.I.).

* Corresponding author. E-mail: Pierre.Hilson@versailles.inra.fr.

The author responsible for distribution of materials integral to the findings present in the Instruction for Authors (www.plantphysiol.org): Pierre Hilson (Pierre.Hilson@versailles.inra.fr).

[W] The online version of this article contains Web-only data.

One-sentence summary

The GLV1-secreted peptide promotes hypocotyl elongation in Arabidopsis and its action requires processing of the precursor by two endoproteases. One of them, SBT6.1, is in turn regulated by a protease inhibitor, suggesting that a complex network controls the action of apoplastic peptide signaling.

Abstract

The *GOLVEN* (*GLV*) gene family encodes small secreted peptides involved in important plant developmental programs. Little is known about the factors required for the production of the mature bioactive GLV peptides. Through a genetic suppressor screen in Arabidopsis (*Arabidopsis thaliana*), we identified two related subtilase genes, *AtSBT6.1* and *AtSBT6.2*, necessary for GLV1 activity. Root and hypocotyl *GLV1* overexpression phenotypes were suppressed by mutations in either subtilase gene. Synthetic GLV-derived peptides were cleaved *in vitro* by the affinity-purified SBT6.1 catalytic enzyme, confirming that the GLV1 precursor is a direct subtilase substrate. Furthermore, the protease inhibitor Serpin1 bound to SBT6.1 and inhibited the cleavage of GLV1 precursors by the protease. *GLV1* and its homolog *GLV2* are expressed in the outer cell layers of the hypocotyl, preferentially in regions of rapid cell elongation. In agreement with the role of the SBT6 subtilases in GLV precursor processing, both null mutants for *sbt6.1* and *sbt6.2* and the *Serpin1* overexpression plants had shorter hypocotyls. The biosynthesis of the GLV signaling peptides requires subtilase activity and may be regulated by specific protease inhibitors. Our data fits with a model in which the GLV1 signaling pathway participates in the regulation of hypocotyl cell elongation, is controlled by SBT6 subtilases, and modulated locally by the Serpin1 protease inhibitor.

Phytohormones are generally considered as the major players in plant intercellular signaling. However, secreted peptides are now also recognized as important molecules in cell-to-cell communication because of their involvement in key developmental processes such as meristem maintenance, organ abscission, cell elongation, cell proliferation and differentiation, gravitropism, and defense (Murphy et al., 2012; Stahl and Simon, 2012). In the complete genome sequence of Arabidopsis (*Arabidopsis thaliana*), more than 1,000 genes have been found that encode putative secreted peptides with a potential signaling function (Lease and Walker, 2006; Ohyama et al., 2008), but the molecular mechanisms that control the production and perception of these peptides have so far only been studied for a few of these genes.

Recently, a novel family of genes has been identified that encodes small secretory peptides designated GOLVEN (GLV), root meristem growth factors (RGF) or CLE-like (CLEL). For clarity, hereafter, these peptides will be referred according to the GLV nomenclature (Fernandez et al., 2013a). The family consists of 11 members that are expressed during different developmental stages and in diverse plant tissues. Particular members show highly specific transcription patterns, usually restricted to a few cell types only (Fernandez et al., 2013b). Some are involved in the control of root meristem maintenance (Matsuzaki et al., 2010), auxin carrier turnover during gravitropic responses (Whitford et al., 2012), root hair formation and lateral root development (Meng et al., 2012; Fernandez et al., 2013b). Specifically, the GLV1 signal modulates auxin gradients in Arabidopsis hypocotyls. Up- or down-regulation of the *GLV1* gene hampers the lateral redistribution of auxin upon gravistimulation of the hypocotyl and inhibits its gravitropic response (Whitford et al., 2012).

Peptides secreted by multicellular eukaryotes are generally synthesized as larger precursor proteins that are biologically inactive and undergo several proteolytic steps, including removal of the signal peptide sequence and subsequent cleavage. In plants, only two enzymes have been shown to process preproproteins into mature signaling peptides. In fact, only a few natural plant protease substrates have been described at all (Tsiatsiani et al., 2012). Finally, additional post-translational modifications are often required to achieve full biological functionality (Matsubayashi, 2011).

In the case of the GLV family, the predicted proteins consist of a central variable region that links two conserved domains: (i) an N-terminal domain coding for a signal peptide that

targets the precursor to the secretory pathway, probably cleaved off by signal peptide peptidases and (ii) a C-terminal domain, designated the GLV motif, that codes for the bioactive mature peptide. For example, the mature bioactive GLV1 peptide consists of a 14-amino-acid sequence derived from the 86-amino acid precursor (Whitford et al., 2012). Thus, proteolytic processing steps are needed to remove portions of the precursor polypeptides leading to the secretion of mature GLV secreted peptides.

GLV proproteins carry sites in their variable region that may be targeted by subtilisin-like serine proteases, also known as subtilases that cleave peptide bonds at or near di-basic residues (Schaller et al., 2012). Compared to other eukaryotic taxons, most subtilase subgroups are underrepresented in plants, whereas those related to pyrolysin have expanded up to 56 members in Arabidopsis, suggesting that these proteases may have evolved with novel target repertoires (Rautengarten et al., 2005). As some subtilases have already been shown to process secreted peptides, they could be involved in the production of GLV signals (Liu et al., 2007; Srivastava et al., 2008, 2009). For example, the Arabidopsis SUBTILASE1.1 (SBT1.1) is required for the processing of the PHYTOSULFOKINE4 (PSK4) propeptide (Srivastava et al., 2008) and SBT6.1 is involved in the maturation of the RAPID ALKALINIZATION FACTOR23 (RALF23) (Srivastava et al., 2009).

Reasoning that some of the Arabidopsis subtilases might be implicated in GLV protein maturation, we investigated their possible requirement for bioactive GLV peptide production. Through a genetic suppressor screen, we identified two subtilases necessary for the GLV1 signal activity and found that they participate in the control of organ growth by modulating cell expansion. The subtilase action on the GLV signaling pathway was confirmed by the characterization of the subtilase biochemical activity on GLV precursors.

RESULTS

Specific subtilases are necessary for GLV1 peptide signaling

Gain-of-function *GLV1* seedlings have an agravitropic curly root when grown on an inclined agar surface (Whitford et al., 2012). Therefore, if a subtilase was responsible for the processing the GLV1 propeptide into its bioactive form, the agravitropic gain-of-function phenotype should be suppressed by a mutation in the corresponding *SBT* gene. Based on this

assumption, we transformed the *GLV1* gene under the control of the *CaMV 35S* promoter into 74 Arabidopsis T-DNA insertion lines, in which 55 of the 56 identified subtilase genes had been mutated (Supplemental Table S1) (Rautengarten et al., 2005). The resulting T1 plants were grown on slanted plates and their root phenotype was scored. Three mutant alleles representing two genes, namely *SBT6.1* (site-1 peptidase or AtS1P; MEROPS ID S08.063; At5g19660) and *SBT6.2* (tripeptidyl-peptidase II; MEROPS ID S08.090; At4g20850), suppressed the agravitropic root phenotype caused by the *GLV1* gain-of-function (Fig. 1A). These subtilase mutants had been identified as SALK_111474 (hereafter designated *sbt6.1-1*), SALK_020530 (*sbt6.1-2*), and SALK_085776 (*sbt6.2*), each carrying a T-DNA insert in an exon (Rautengarten et al., 2005) (Supplemental Table S1).

Several independent *GLV1*-overexpressing ($GLV1^{OE}$) homozygous lines were obtained for each of the three *sbt* mutant genotypes and those with a high level of *GLV1* transcripts were selected for further study. In all cases, quantitative reverse-transcription polymerase chain reaction (qRT-PCR) analysis confirmed that the suppression was linked to the lack of expression of the *SBT6.1* or *SBT6.2* subtilase gene, and that in selected transformed lines the *GLV1* gene was overexpressed at levels that cause agravitropic root growth in wild-type (WT) plants (Fig. 1B; Supplemental Table S2).

Root growth phenotypes were quantified by measurement of the gravitropic index (GI), which is the ratio between the primary root length and the linear distance separating the collet from the root tip (Fig. 1C) (Grabov et al., 2005). Comparative analysis confirmed that the strong gravitropic defect of $GLV1^{OE}$ roots is suppressed in the *sbt6.1-1* and *sbt6.2* mutants, but also revealed that the *SBT6* loss-of-function alone resulted in a phenotype opposite to that of the *GLV1* gain-of-function: the single *sbt6.1-1* and *sbt6.2* lines and the double loss-of-function *sbt6.1-1 sbt6.2* line had a GI that was higher than that of the WT. These results suggest that SBT6.1 and SBT6.2 are necessary for the processing of the GLV1 and possibly, other GLV precursors involved in root gravitropic responses (Whitford et al., 2012).

GLV and SBT6 functions interact to control hypocotyl elongation

Whereas our screen was based on a root phenotype resulting from overexpression, the native *GLV1* gene is not transcribed in the root, but in the aerial part of Arabidopsis plants,

including the growing hypocotyl, together with its close homolog *GLV2* (Whitford et al., 2012; Fernandez et al., 2013b) (Fig. 2). *SBT6.1* is also transcribed in the growing hypocotyl (Fig. 2) (Liu et al., 2007). Based on their common expression domain, we speculated that these three genes are involved in hypocotyl development. Compared to the WT, the elongation of $GLV1^{OE}$ and $GLV2^{OE}$ hypocotyls grown in the dark was faster, whereas that of *sbt6.1-1*, *sbt6.2*, and *sbt6.1-1 sbt6.2* was slower (Fig. 3, A and B). Furthermore, the increase in hypocotyl size induced by the overexpression of *GLV1* or *GLV2* was suppressed in the *sbt6* loss-of-function mutants (Fig. 3B). The difference in hypocotyl growth between genotypes was not the indirect consequence of early developmental delays, because all lines germinated simultaneously and their hypocotyl length was undistinguishable at the beginning of the experiment (Fig. 3A). Thus, these observations indicate that the subtilases positively control hypocotyl elongation, possibly through processing of GLV peptides.

Single loss-of-function *amiRglv1* (artificial microRNA interference knockdown) and *glv2-1* (T-DNA knockout) lines and the double *amiRglv1 glv2-1* line had no significant defect in hypocotyl growth, probably because of the action of partially redundant *GLV* genes. For example, the *GLV10* transcript was detected in the growing hypocotyl (Fernandez et al., 2013b). Alternatively, additional non-GLV signaling peptide precursors that positively regulate hypocotyl elongation may also need to be processed by the subtilases to become active.

To better understand the mode of action of the GLV signal at the cellular level, we measured the length of hypocotyl epidermal cells, where *GLV1* and *GLV2* are primarily transcribed, in GLV^{OE} and *SBT6* mutant lines (Supplemental Fig. S1) (Whitford et al., 2012). The cells were longer in plants overproducing the GLV1 or GLV2 peptide and shorter in the *sbt6.1-1* and *sbt6.2* lines as compared to those of the WT (Fig. 3C; Supplemental Fig. S1). Our results imply that GLV signaling positively regulates hypocotyl growth by promoting cell elongation.

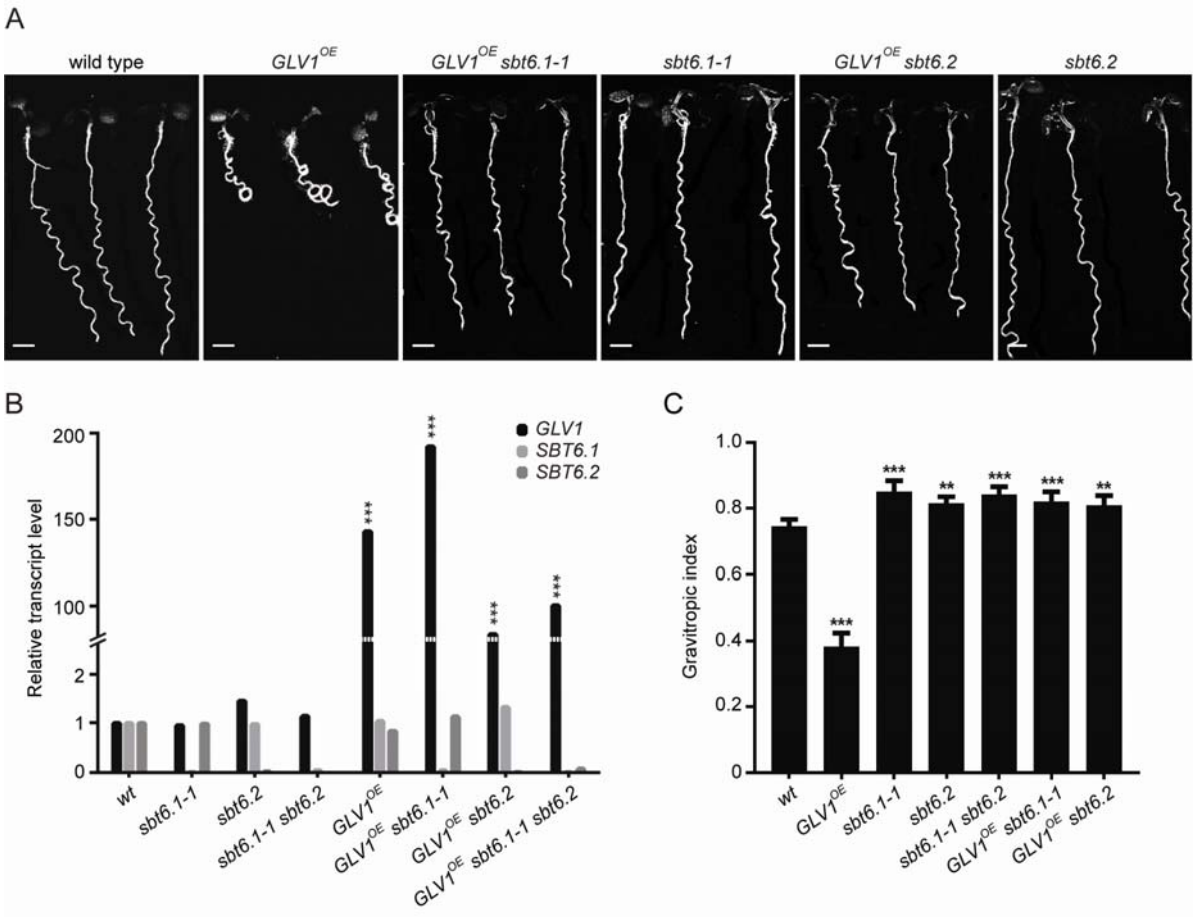


Fig. 1. Suppression of the *GLV1* overexpression curly root phenotype in subtilase mutants.

A, Seedlings grown on inclined agar plates for 7 days after germination (dag). Scale bars, 2 mm. B, Relative transcript levels compared to WT as measured by qRT-PCR analysis (mean transcript level ± confidence interval [CI]; one-way ANOVA). C, Gravitropic index (mean GI index ± CI compared to WT; one-way ANOVA, n = 18-39). Error bars represent the 95% confidence interval. Asterisks mark significant differences: **P*<0.05; ***P*<0.005; ****P*< 0.001.

The GLV1 peptide promotes hypocotyl elongation and rescues the *sbt6* phenotype

The bioactive peptide is encoded in a conserved C-terminal motif of the GLV1 precursor (Whitford et al., 2012). To confirm that the GLV1-induced hypocotyl growth can be attributed to that domain, we treated Arabidopsis seedlings with a synthetic peptide (GLV1p) similar to the mature native signal, DY(SO_3H)PQPHRKPPIHNE, with Y(SO_3H) a sulfated tyrosine. Hypocotyls of plants incubated for 5 days in liquid medium supplemented with 1 µM GLV1p were longer than those of control plants (Fig. 3D).

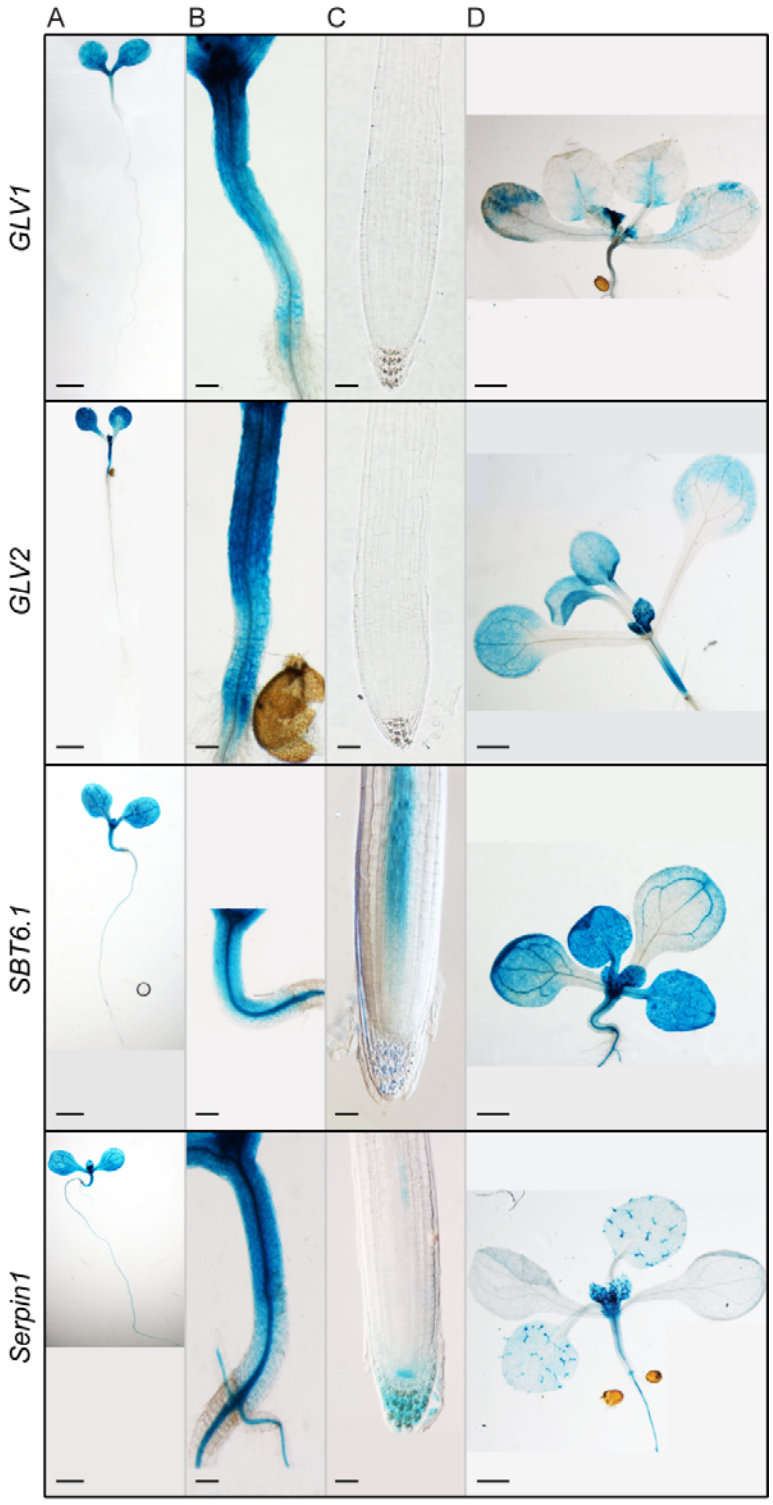


Fig. 2. Transcriptional activity of *GLV1*, *GLV2*, *SBT6.1* and *Serpin1*.
A-C, Young seedling, hypocotyl, and root tip (5 days after germination [dag]). D, Cotyledons and first leaves (10 dag). Plants were transformed with the corresponding *promoter:GUS* transgene. [Scale bars: 1 mm (A and D), 100 µm (B), 50 µm (C)].

If SBT6 subtilases were involved in the processing of the GLV1 precursor, then the *sbt* loss-of-function mutants should still respond to the addition of the GLV1 mature peptide. To test this assumption, we measured the effect of GLV1p on *sbt6.1* and *sbt6.2* plants: the hypocotyl length of the mutants treated with GLV1p was longer than that of the untreated counterparts, reaching a size undistinguishable from that of untreated WT control plants (Fig. 3D). This peptide effect together with the failure of $GLV1^{OE}$ plants to produce longer hypocotyls in the *sbt6.1* and *sbt6.2* backgrounds indicate that the subtilases act upstream of the GLV signal perception.

GLV precursors are proteolytically cleaved by SBT6.1

To investigate whether the SBT6 proteolytic activity might be involved directly in the processing of the GLV1 precursor protein, we overproduced the myc epitope-tagged SBT6.1 protein in Arabidopsis (Srivastava et al., 2009) and affinity purified the subtilase from whole plants germinated and grown in liquid medium. Two canonical subtilase recognition sequences were identified within the variable region of the GLV1 precursor (Fig. 4A). The SBT6.1 enzyme bound to anti-myc beads was incubated with synthetic propeptides that corresponded to these regions (Fig. 4A; Supplemental Table S3). As a control, the subtilase proteolytic activity was confirmed with a synthetic RALF23 propeptide that is a known SBT6.1 substrate (Fig. 4B; Supplemental Table S3) (Srivastava et al., 2009). The enzymatic digestion products were analyzed by matrix-assisted laser desorption/ionisation-time of flight mass spectrometry (MALDI-TOF MS). The tested GLV1 propeptide-derived products revealed two main SBT6.1 cleavage sites after the RRLR and RRRAL sequences, and a minor one ending with RRRA (Fig. 4, A and C; Supplemental Fig. S3; Supplemental Table S3). As the latter sequence did not contain the leucine residue part of the SBT6.1 recognition site defined previously as RXLX or RXXL (Liu et al., 2007), its presence may not be a strict requirement for SBT6.1 proteolysis.

In summary, SBT6.1 cleaves the GLV1 precursor protein in vitro. Furthermore, null mutations in *SBT6.1*, and its closest homolog *SBT6.2*, suppress *GLV1* gain-of-function phenotypes. The *in vivo* and *in vitro* results suggest that the SBT6 activity is needed for maturation and activation of the GLV1 propeptide.

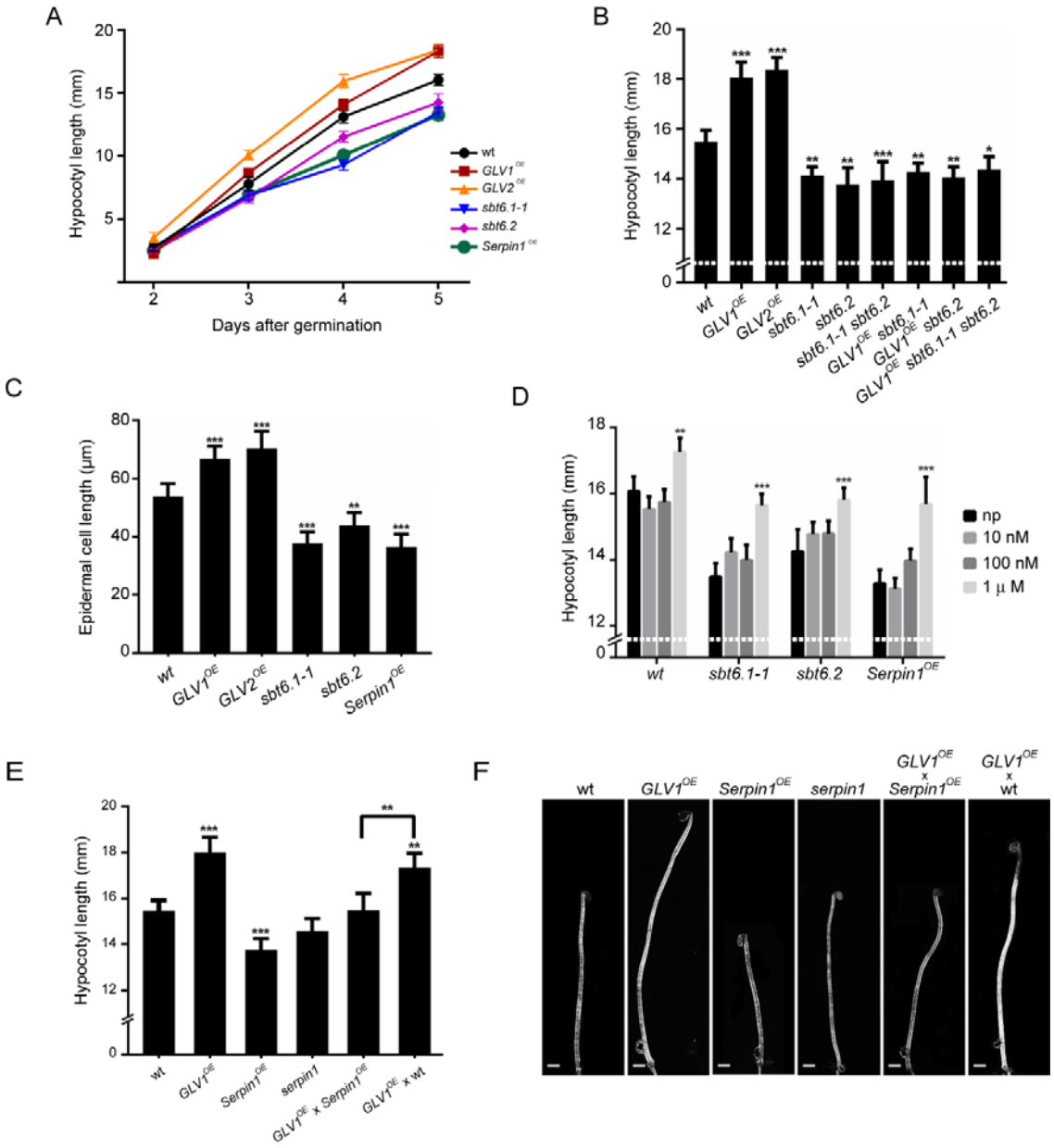


Fig. 3. Hypocotyl elongation phenotypes.

A, Kinetics of etiolated hypocotyl growth (mean hypocotyl length in mm ± CI; two-way ANOVA; n = 32-80). B, Hypocotyl length 5 dag (mean hypocotyl length in mm ± CI compared to WT; one-way ANOVA; n = 32-65). C, Hypocotyl epidermal cell length (mean of three most elongated cells from each seedling in µm, one-way ANOVA; n = 15). D, Hypocotyl length (in mm) upon GLV1 peptide treatment. Peptide treatments at different concentrations were compared to mock-treated (without peptide; np) plants of the same genotype 5 dag (two-way ANOVA; n = 52-100). Error bars represent the 95% confidence interval. Asterisks mark significant differences: **P*<0.05; ***P*<0.005; ****P*< 0.001. E, Hypocotyl length at 5 dag (mean hypocotyl length in mm ± CI compared to WT; one-way ANOVA; n = 20-61). Hemizygous F1 plants were measured to assess the interaction between the *GLV1* and *AtSerpin1* gain-of-

function, in comparison to F1 plants resulting from a cross between *GLV1*^{OE} and WT plants. F, Representative hypocotyl length. [Scale bars: 1 mm]. Error bars represent the 95% confidence interval. Asterisks mark significant differences: *P<0.05; **P<0.005; ***P<0.001.

SBT6.1 associates with the Serpin1 protease inhibitor

The activity of proteases involved in the control of developmental processes has to be tightly regulated (Turk, 2006). In separate tandem affinity purification (TAP) experiments, we noted that SBT6.1 occurred *in vivo* in protein complexes formed with the Serpin1 protease inhibitor (MEROPS ID I04.087; also referred to as AtSerpin1) (Supplemental Table S4).

The close association of SBT6.1 and Serpin1 was confirmed by bimolecular fluorescence complementation (BiFC) analysis. Both proteins were transiently coproduced as translational fusions with the truncated halves of the enhanced green fluorescent protein (EGFP) in epidermal cells of *Agrobacterium tumefaciens*-transfected *Nicotiana benthamiana*. SBT6.1 was fused to the N terminus of EGFP (nGFP) and combined with the Serpin1 protein that had been fused to the C terminus of EGFP (cGFP), and vice versa. In all cases, the EGFP fragments were fused at the C end of the tested interactors. In both configurations, AtSBT6.1 and Serpin1 interactions resulted in a strong apoplastic signal (Fig. 5). As negative controls, single constructs (either Serpin1 or SBT6.1 fused to nGFP or cGFP) were infiltrated, but without any detectable signal (data not shown). We should mention, however, that this approach has certain limitations such as the delay between the time when the fusion proteins interact with each other and the time when the complex becomes fluorescent. Moreover there seems to be a tendency of the two GFP halves fragments to attach to each other (Kerppola , 2010). Therefore, it is necessary to further study the biological interaction between SBT6.1 and Serpin1 using more powerful tools such as fluorescence resonance energy transfer (FRET) and yeast two hybrids. An additional control including transformation of the two GFP halves fragments without any protein fusion should be also included in the BiFC experiment.

To investigate whether the association of SBT6.1 with a protease inhibitor negatively regulated the subtilase proteolytic activity, we measured the cleavage of the GLV1 precursor sequences by the SBT6.1 enzyme in the presence of Serpin1. As expected, no digested peptide could be detected by MALDI-TOF when Serpin1 was added to the purified SBT6.1 protein (Fig. 4, D and E). These results demonstrated that Serpin1 inhibits the SBT6.1 activity *in vitro*.

Serpin1 overexpression suppresses GLV-dependent hypocotyl elongation

Our biochemical analysis pointed toward a potential role of Serpin1 in GLV-dependent regulation of hypocotyl elongation through the control of the SBT6.1 proteolytic activity. In agreement with this model, the *Serpin1* promoter was active in the hypocotyls, as well as other plant parts (Fig. 2; Supplemental Fig. S2).

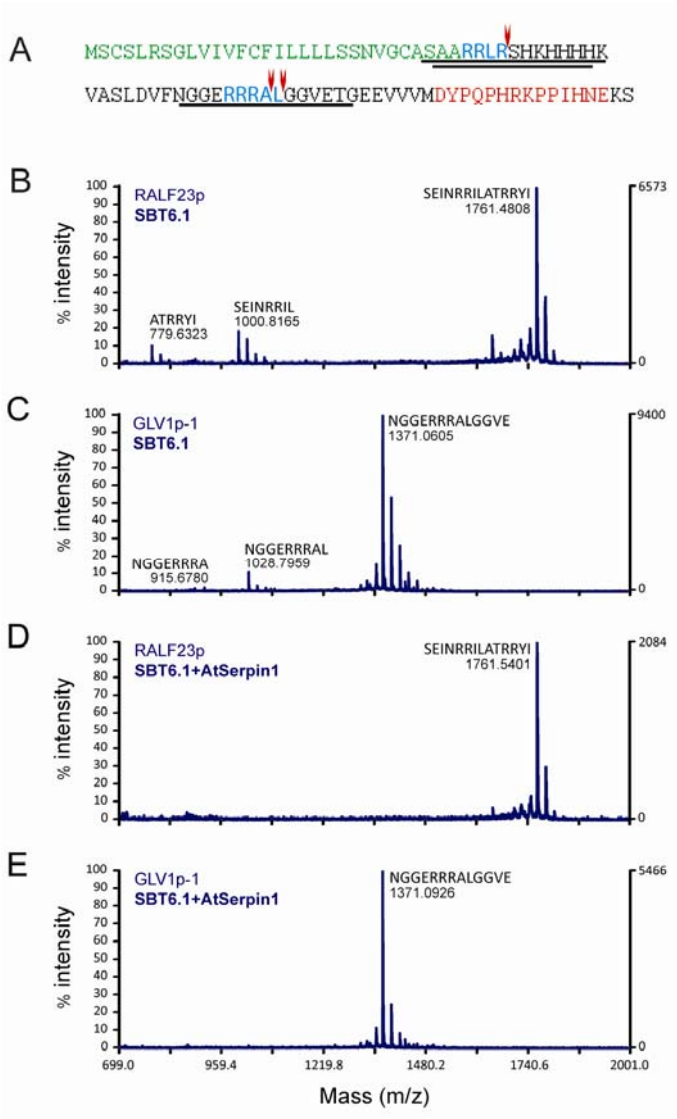


Fig. 4. *In vitro* proteolytic activity of SBT6.1. A, GLV1 propeptide sequence. Signal peptide, green; subtilase canonical cleavage site, blue; GLV motif, red; observed cleavage sites, red arrows; tested synthetic peptides, underlined. Synthetic peptide and fragments after 1 h of incubation with SBT6.1 for RALF23 (B and D) and GLV1 (C and E) fragments, and with (D and E) or without (B and C) AtSerpin1.

The hypocotyls of Serpin1–overexpressing (*Serpin1*^{OE}) plants were shorter and their epidermal cells smaller than in WT plants (Fig. 3, C-E). These phenotypes were reminiscent of those observed in *sbt6.1-1* and *sbt6.2* loss-of-function mutants (compare Fig. 3, E and F with Fig.

3, B and D). Furthermore, the long-hypocotyl phenotype associated with the *GLV1* gain-of-function was suppressed in *Serpin1*^{OE} seedlings (Fig. 3, E and F). Finally, *Serpin1*^{OE} seedlings treated with the bioactive synthetic GLV1p had longer hypocotyls than the untreated seedlings, with a response similar to that of *sbt6.1-1* and *sbt6.2* loss-of-function mutants, thereby confirming that the SBT6 and Serpin1 activities are involved in the production of the GLV signal (Fig. 3D).

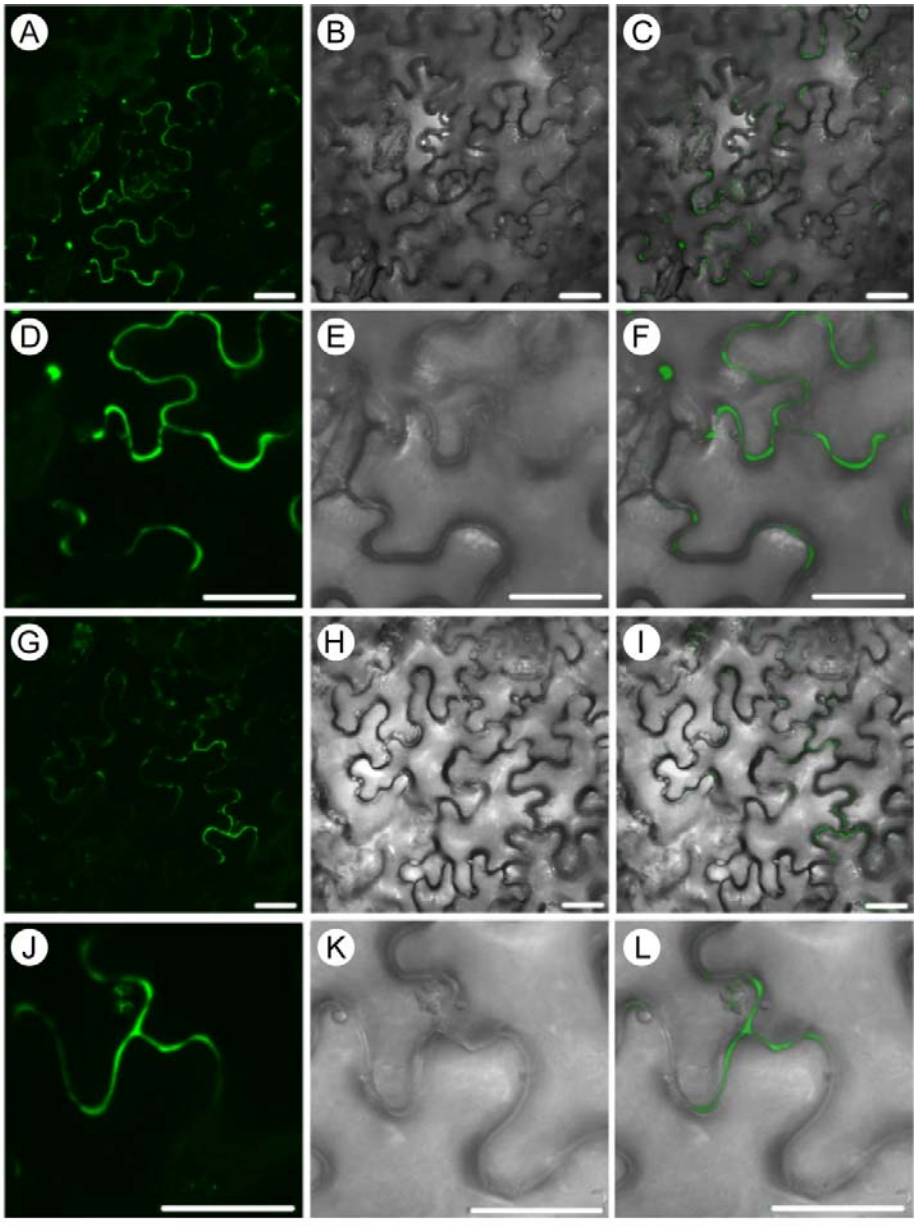


Fig. 5. BIFC interaction between Serpin1 and SBT6.1.

A-F, Interaction between SBT6.1-nGFP and Serpin1-cGFP. G-L, Interaction between SBT6.1-cGFP and Serpin1-nGFP in the leaf epidermis of *Nicotiana benthamiana*. GFP fluorescence (A, D, G, and J), Nomarski differential interference contrast (DIC) (B, E, H, and K), and GFP/DIC overlapping images (C, F, I, and L*)*. All images resulted from stacked confocal sections. [Scale bars: 25 µm.]

DISCUSSION

The catalytic processing of subtilases is required for GLV peptide production

Our initial suppressor screen based on *GLV1* overexpression root phenotypes and the subsequent analysis of related hypocotyl growth phenotypes revealed that the genes coding for SBT6.1 and SBT6.2 are necessary for the maturation and activation of the GLV1 peptide. These two proteins are most closely related to each other in the subtilase phylogenetic tree and may, therefore, have similar activities (Rautengarten et al., 2005). However, the *sbt6.1-1 sbt2* double mutant did not exhibit an additive phenotype, suggesting that the two subtilases act instead at successive stages during GLV1 maturation.

In vitro protease assays showed that the plant-purified SBT6.1 enzyme cleaves GLV1 precursor peptides at sites reminiscent of the canonical recognition sequences for subtilases, RXXL and RXLX (Schaller et al., 2012). As 10 out of 11 GLV precursors carry at least one of these sites (Table S5), SBT6.1 may cleave multiple members of the GLV family. The majority of GLV peptides are expressed in root tissues, some of which are involved in the root gravitropic response (Whitford et al., 2012; Fernandez et al., 2013b). The fact that s*bt6.1* and *sbt6.2* loss-of-function mutants have a higher gravitropic index than the WT hints at the involvement of *SBT6.1* and *SBT6.2* in the maturation of GLV precursors produced in the primary Arabidopsis root, other than GLV1.

Finally, the SBT6.1 cleavage sites identified in the GLV1 precursor sequence are not sufficient to produce the mature peptide detected in plant tissues (Matsuzaki et al., 2010). The additional processing steps are probably catalyzed by other proteases, such as SBT6.2, homolog of the mammalian tripeptidyl peptidase TPPII, the proteolytic activity of which is quenched by TPPII-specific inhibitors (Book et al., 2005). Carboxypeptidases may also be involved, such as SUPPRESSOR OF LLP1 1 (SOL1), which cleaves C-terminal lysine and arginine residues off the end of CLE peptides and is required for CLE19 signaling (Casamitjana-Martínez et al., 2003; Tamaki et al., 2013).

Protease inhibitors as peptide signaling modulators

Among the 68 peptidase inhibitor families, the serpins are one of the two largest groups that can be found in all kingdoms. Serpins are suicide inhibitors that form irreversible covalent complexes with their targets (Huntington et al., 2000). Their function in plants remains poorly understood (Fluhr et al., 2012). Thus far, the Arabidopsis *Serpin1* has been shown to be involved in the inhibition of the Arabidopsis metacaspase 9 and the vacuolar protease RESPONSIVE-TO-DESICCATION21 (RD21) (Vercammen et al., 2006; Lampl et al., 2010, 2013).

Our data indicate that SBT6.1 is also under the control of the Serpin1 protease inhibitor: tandem affinity purification and BiFC analyses have shown that both proteins interact and that the proteolytic activity of the subtilase is inhibited by Serpin1 in vitro. Furthermore, *Serpin1* overexpression phenocopies *sbt6.1* and *sbt6.2* null mutants and suppresses *GLV1* gain-of-function. These observations can be summarized in a model, in which hypocotyl cell elongation is positively regulated by the GLV1 peptide and its production is catalyzed by SBT6.1 that is itself inhibited by Serpin1 (Fig. 6).

Serpins carry a reactive center loop (RCL), including a protease target sequence as bait. Upon cleavage of this target sequence, the RCL undergoes an irreversible conformational change that locks and inactivates the protease (Lampl et al., 2010). Serpin1 contains the canonical cleavage site RGLL that makes it a potential target for SBT6.1, in agreement with the protein interaction and biochemical analyses of the SBT6.1 proteolytic activity.

The SBT6.1 protein has been located in the Golgi apparatus (Liu et al., 2007) and Serpin1 has been detected in the cytosol, Golgi bodies, endoplasmic reticulum and apoplast (Vercammen et al., 2006; Lampl et al., 2013). Therefore, SBT6.1 and Serpin1 can interact to regulate GLV-dependent hypocotyl growth. Nevertheless, because SBT6.1 is not the sole target of Serpin1, we cannot exclude that the inhibition effect on the hypocotyl growth may also be partly relayed through inactivation of other proteases.

The 11 Arabidopsis *GLV* genes share sequence similarity, but they are expressed specifically in different tissues (Whitford et al., 2012; Fernandez et al., 2013b), in contrast to the relatively broad expression patterns of the *SBT6.1* and *Serpin1* genes in Arabidopsis. Therefore, they may be involved in the coregulation of GLV functions in various tissues, including, but not

exclusively, in the hypocotyl. For example, seven out of 11 *GLV* genes are transcribed in the root tip (Fernandez et al., 2013b) where *SBT6.*1 and *Serpin1* are expressed as well (Supplemental Fig. S1). Whereas *Serpin1* was transcribed in all tested organs (Ahn et al., 2009), its expression is seemingly not uniform across all cell types (Supplemental Fig. S2). Hence, we postulate that, by limiting the SBT6.1 activity, Serpin1 contributes to the fine spatial regulation of the GLV peptide biosynthesis.

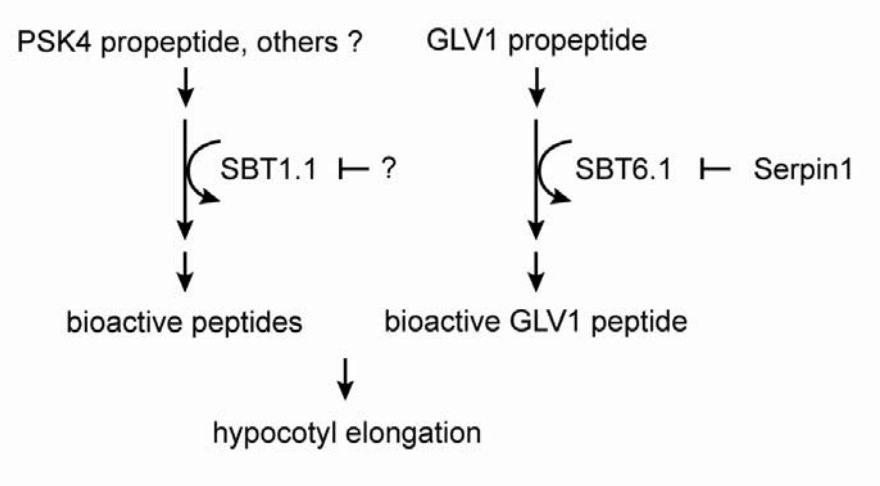


Fig. 6. Model for *GLV*-dependent hypocotyl elongation.

GLV peptides control hypocotyl elongation together with other secreted peptides

Our experimental results demonstrate that GLV signals promote cell elongation in the growing hypocotyl: (i) overexpression of the *GLV1* and *GLV2* genes, normally transcribed in the outer cell layers of the hypocotyl (Whitford et al., 2012), results in longer epidermal cells; (ii) application of the bioactive synthetic GLV1 peptide increases the hypocotyl length; and (iii) null mutations in genes coding for subtilases necessary for the proteolytic processing of the GLV1 precursor cause an opposite short-hypocotyl phenotype. These observations confirm that *GLV1* and *GLV2* play a positive role in the cell expansion regulation, as already implied by their requirement for the gravitropic response of reoriented hypocotyls (Whitford et al., 2012).

Whereas highly significant, the differences observed in hypocotyl lengths are limited to 10-20% gain or loss when compared to those of WT, indicating that other signals also take part in the control of the hypocotyl growth. Other secreted peptides have been shown to promote

hypocotyl cell expansion, including PSK-α (Fig. 6) and PLANT PEPTIDE CONTAINING SULFATED TYROSINE 1 (PSY1) (Amano et al., 2007; Stührwohldt et al., 2011; Hartman et al., 2013). Conspicuously, mature GLV, PSK and PSY peptides all carry a sulfated-tyrosine residue that is important for bioactivity and results from the activity of the tyrosylprotein sulfotransferase (TPST) (Matsybayashi, 2011). They may also share other processing enzymes, including subtilases. Yet, a possible crosstalk between the peptide signaling pathways driving cell expansion remains to be elucidated, as well as their connection with classical hormonal growth control.

MATERIALS AND METHODS

Growth conditions

Unless otherwise specified, seeds were surface-sterilized and sown on half strength Murashige and Skoog (MS) medium (Duchefa Biochemie B.V.) complemented with 1% (w/v) agarose and 1.5% (w/v) sucrose at pH 5.8, and stratified for at least 2 days at 4°C. Seedlings were germinated in illuminated growth chambers under a 16-h light/8-h dark cycle (100 µmol m^{-2} s^{-1}) at 21°C. For root growth analysis, plates were slanted at a 45° angle with respect to the gravity vector for 7 days. For hypocotyl length measurements, seeds were surface-sterilized, stratified at 4°C for at least 2 days in liquid MS media, exposed to the light for 6 h, and then transferred to darkness for 2 to 5 days under continuous rotation. Imaged root and hypocotyl features were measured with the ImageJ software (http://rsbweb.nih.gov/ij/).

Recombinant DNA constructs and Arabidopsis lines

The *GLV1*^{OE}, *GLV2*^{OE}, *GLV1pro:GUS-GFP*, *GLV1pro:NLS-GFP-GFP*, *GLV2pro:GUS-GFP* and *GLV2pro:NLS-GFP-GFP* lines have been described previously (Fernandez et al., 2013b). The subtilase mutant collection was a gift from Dr. Thomas Altmann (Institut für Biochemie und Biologie, Genetik, Universität Potsdam, Golm, Germany). The presence of T-DNA inserts in the *Arabidopsis SBT* genes was confirmed by PCR analysis with gene-specific primers and a left border T-DNA primer (for all primer sequences, see Supplemental Table S6). The *GLV1*^{OE} *sbt* mutant lines were produced by introducing the *35S:GLV1* construct (carrying either the kanamycin or Basta resistance gene) into the *sbt* mutant lines via floral dip (Clough and Bent, 1998). At least five independent *GLV1* gain-of-function lines were analyzed per transformed mutant lines. Plant DNA was isolated and analyzed by PCR. The double *sbt6.1-1 sbt-6.2* knockout mutant lines, with or without the *GLV1*^{OE} transgene, were obtained through crosses and genotyped at the F2 generation (for primer sequences, see Supplemental Table S6).

To generate *GLV1*^{OE} and *Serpin1*^{OE} cassettes the full-length coding sequences (CDS) of the genes were amplified by PCR from first-strand cDNA of *Arabidopsis thaliana* L. Heyhn. (accession Columbia [Col-0]) with gene-specific primers extended with either the *att*B1 or *att*B2

sites for the Gateway recombinational cloning. The resulting PCR fragments were captured by BP clonase reaction in an entry clone derived from pDONR221. Overexpression constructs were obtained by LR recombination between the entry clones and the destination vector pK7GW2 or pB7GW2 (3). PCR reactions were run with High Fidelity Platinum Taq DNA Polymerase (Invitrogen).

For *Serpin1pro:GUS-GFP,* the promoter sequence (~1500 bp upstream of the start codon) was amplified by PCR from the Arabidopsis Col-0 genomic DNA with Gateway-compatible primers (Supplemental Table S6). The promoter amplicon was cloned into pBGWFS7 (Karimi et al., 2002) generating pBGWFS7PAtSRP1 that codes for a transcriptional fusion with a *GFP:GUS* translational fusion gene. The BiFC expression clones (p35S:ORF:nGFP and p35S:ORF:cGFP) were generated in the pK7m34GW destination vector (http://www.psb.ugent.be/gateway/index.php) (Boruc et al., 2010). In all cases, the EGFP fragments were fused at the C end of the tested interactors.

Gene expression analysis

Total RNA from 3-week-old leaves was isolated using TRIzol reagent (Invitrogen), followed by treatment with RNase-free DNase I (Qiagen) according to the manufacturer's instructions. The cDNA was prepared with the iScript™cDNA Synthesis Kit (Bio-Rad) from 1 μg of total RNA. For quantitative RT-PCR, 1:10 dilutions of total cDNA were used (all the primers are listed in Supplemental Table S6).

Peptide treatments

Seedlings were germinated, grown in liquid medium supplemented or not with the synthetic GLV1p, and incubated for 6 h in the light in rotating six-well plates at 21°C, and then for 5 days in the dark under continuous rotation. The peptide was synthesized in house as previously described (Whitford et al., 2012) and dissolved in sterile sodium phosphate buffer (50 mM, pH 6.0).

Purification of the SBT6.1 enzyme from plant tissues

The myc-epitope-tagged SBT6.1 protein was affinity purified from an overexpression line previously described (Srivastava et al., 2009). Plantlets for protein extraction were grown in liquid half-strength MS medium with orbital shaking at 130 rpm (Innova^{TM} 2300, New Brunswick Scientific). Fifty grams of 2-week-old Arabidopsis seedlings grown in liquid half-strength MS medium were ground in liquid nitrogen and suspended in ice-cold extraction buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 0.1% Nonidet P-40, and 10% ethylene glycol) with an ultra-Turrax mixer. The supernatant was centrifuged twice at 18,000*g*. A 100-µL volume of anti-c-myc Agarose Affinity Gel (Sigma-Aldrich) was added to the filtered lysate and incubated for 2 h at 4°C with continuous rotation. The SBT6.1 enzyme bound to the agarose affinity gel was recovered by centrifugation at 1,500*g* for 4 min at 4°C and washed 10 times thoroughly with washing buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl) through a polyprep chromatography column (Bio-Rad). The final product was resuspended in 100 µL of 25 mM 2-(*N*-morpholino)-ethanesulfonic acid (MES)-sodium acetate buffer (pH 6.2). The bead-bound protein concentration was measured with the protein assay (Bio-Rad). All purified products were resolved on 10% SDS-PAGE gel and visualized either by Coomasie brilliant blue staining or Western blots with the 9E10 monocolonal anti-myc antibody (Santa Cruz Biotechnology). For negative controls, nontransgenic plants were purified in parallel.

Peptide assay for SBT6.1 activity in vitro and its inhibition by Serpin1

For protease activity assays with RALF23 and GLV1 propeptides, 19 µL of bead-bound affinity-purified myc-tagged SBT6.1 was mixed with 1 µL of a 500-mM peptide solution in 25 mM MES-sodium acetate buffer (pH 6.2), supplemented with 2.5 mM calcium chloride, to obtain a final peptide concentration of 25 µM. Standard enzymatic reactions were incubated at 32°C for 1 h.

Serpin1 was purified from *Escherichia coli* cultures as described (Vercammen et al., 2006). Of bead-bound affinity-purified myc-tagged SBT6.1, 18 µL was mixed with Serpin1 in phosphate-buffered saline/glycerol (50:50) to a final concentration of 0.5 mg/mL and a total volume of 19 µL. Beads loaded with myc-tagged SBT6.1, but without Serpin1, were used as

positive controls. The beads were incubated for 1 h at 32°C as described above. The SBT6.1 peptide digestion products were analyzed by mass spectrometry with a Voyager DE STR MALDI-TOF spectrometer (Applied Biosystems). The matrix contained 4 to 5 mg α-cyano-4-hydroxycinnamic acid in 1 mL acetonitrile/MilliQ water (50:50) supplemented with 10 mM ammoniumcitrate and 1 µL trifluoroacetic acid. The crude peptide mixture was spotted on the MALDI-TOF plate and analyzed.

Analysis of SBT6.1 and Serpin1 association

In vivo interaction of SBT6.1 with Serpin1 was determined by tandem affinity purification as described (Van Aken et al., 2007).

In brief, Arabidopsis cell suspension cultures were stably transformed by *Agrobacterium tumefaciens*-mediated cocultivation with pKNTAP-Serpin1. The TAP tag consisted of two IgG-binding domains of the *Staphylococcus aureus* protein A (ZZ) and a calmodulin-binding peptide (CBP), separated by a tobacco etch virus (TEV) protease cleavage site (Rigaut et al., 1999). Two-step affinity purification was done as described (Van Leene et al., 2007). To increase the stringency of the data set, proteins commonly contaminating complex extracts were considered as experimental background and systematically subtracted from the lists of copurified proteins (Van Leene et al., 2010).

Histochemical and microscopic analysis

GUS staining was performed as described previously (Beeckman and Engler, 1994). For live-cell imaging, seedlings were mounted in water with or without dye. The adaxial leaf epidermis of transfected *N. benthamiana* leaves were assayed for fluorescence with a confocal microscope LSM5 (Zeiss) equipped with a 40× and 63× water-corrected objectives. GFP fluorescence was imaged with a 488-nm laser excitation. Emission fluorescence was captured in the frame-scanning mode alternating GFP fluorescence via a 500-/550-nm band-pass emission filter. Cell membranes of hypocotyls were counterstained with propidium iodide and imaged with a 543-nm filter and 590- to 620-nm for excitation and detection, respectively.

Transient expression in *N. benthamiana*

Wild-type *N. benthamiana* plants were grown under 14 h of light and 10 h of darkness at 25°C and 70% relative humidity. All BiFC constructs were transferred into the *A. tumefaciens* strain C58C1 harboring the virulence plasmid MP90 (Boruc et al., 2010). The obtained *Agrobacterium* strains were used to infiltrate tobacco leaves, of which the transient expression was assayed. The transformed *Agrobacterium* strain harboring the constructs of interest was grown in 2 mL of yeast extract broth (YEB) supplemented with appropriate antibiotics in a shaking incubator (200 rpm) at 28° C. After 1 day, 100 µL of the liquid culture was transferred to 10 mL YEB supplemented with appropriate antibiotics and grown for 1 additional day. After incubation, the OD_{600} of each culture was measured, the amount of culture needed for OD_{600} = 1.5 was transferred to Eppendorf tubes, and centrifuged at 6800*g* for 5 min. The bacterium pellet was resuspended in 2 mL of the infiltration buffer (10 mM $MgCl_2$, 10 mM MES, and 100 µM acetosyringone) and incubated for 2 h as described above. For coexpression experiments, 0.33 mL of each bacterial culture was mixed with the bacterial culture harboring p19 vector to obtain 1 mL of the inoculum, with each construct adjusted to final OD_{600} value = 0.5. The inoculum was delivered to 3- to 4-week-old *N. benthamiana* leaves by gentle pressure infiltration of the abaxial epidermis with a 1-mL syringe without needle. The infiltrated leaf areas were delimited and labeled with an indelible pen. Plants were further grown under normal growing conditions. Four infiltrated leaf fragments were analyzed per combination in two independent transformation events 2, 3, and 5 days after infiltration. Interactions were scored positive when at least 10 fluorescent cells per leaf segment were observed. Infiltrated leaves were imaged with a LEICA LCS-SL CLSM confocal microscope.

Statistical tests

Means of samples were compared with one-way or two-way analysis of variance (ANOVA) (GraphPad Prism; V6.00, GraphPad Software). Data were pooled from two independent biological replicates unless specified otherwise.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Hypocotyl elongation phenotypes.

Supplemental Figure S2. Transcriptional activity of pSerpin1:GUS.

Supplemental Figure S3. MALDI-TOF spectra for synthetic peptides.

Supplemental Table S1. Subtilase (*sbt*) mutant genotypes of Arabidopsis.

Supplemental Table S2. *GLV1* transcript fold induction in transformed *sbt6* T-DNA mutant lines.

Supplemental Table S3. Propeptides and observed proteolytic products according to m/z for singly charged ions.

Supplemental Table S4. Proteins identified after TAP purification with NTAP-Serpin1 expressed in Arabidopsis cell suspension culture.

Supplemental Table S5. Typical subtilases target sequences in GLVs.

Supplemental Table S6. Primers used to confirm overexpression and knockout lines.

ACKNOWLEDGEMENTS

We thank Carsten Rautengarten, Thomas Altmann, and Stephen Howell for the generous gift of Arabidopsis subtilase mutant and reporter lines. We thank Martine De Cock for help in preparation of the manuscript.

LITERATURE CITED (main text + Suppl. M&M)

Ahn J-W, Atwell BJ, Roberts TH (2009) Serpin genes *AtSRP2* and *AtSRP3* are required for normal growth sensitivity to a DNA alkylating agent in *Arabidopsis*. BMC Plant Biol **9**: 52

Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. Proc Natl Acad Sci USA **104**: 18333-18338

Beeckman T, Engler G (1994) An easy technique for the clearing of histochemically stained plant tissue. Plant Mol Biol Rep **12**: 37-42

Book AJ, Yang P, Scalf M, Smith LM, Vierstra RD (2005) Tripeptidyl peptidase II. An oligomeric protease complex from Arabidopsis. Plant Physiol **138**: 1046-1057

Boruc J, Van den Daele H, Hollunder J, Rombauts S, Mylle E, Hilson P, Inzé D, De Veylder L, Russinova E (2010) Functional modules in the *Arabidopsis* core cell cycle binary protein-protein interaction network. Plant Cell **22**: 1264-1280

Casamitjana-Martínez E, Hofhuis HF, Xu J, Liu C-M, Heidstra R, Scheres B (2003) Root-specific *CLE19* overexpression and the *sol1/2* suppressors implicate a CLV-like pathway in the control of *Arabidopsis* root meristem maintenance. Curr Biol **13**: 1435-1441

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735-743

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, Hilson P (2013a) Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol **161**: 954-970

Fernandez A, Hilson P, Beeckman T (2013b) GOLVEN peptides as important regulatory signalling molecules of plant development. J Exp Bot **64**: 5263-5268

Fluhr R, Lampl N, Roberts TH (2012) Serpin protease inhibitors in plant biology. Physiol Plant **145**: 95-102

Grabov A, Ashley MK, Rigas S, Hatzopoulos P, Dolan L, Vicente-Agullo F (2005) Morphometric analysis of root shape. New Phytol **165**: 641-651

Hartmann J, Stührwohldt N, Dahlke RI, Sauter M (2013) Phytosulfokine control of growth occurs in the epidermis, is likely to be non-cell autonomous and is dependent on brassinosteroids. Plant J **73**: 579-590

Huntington JA, Read RJ, Carrell RW (2000) Structure of a serpin--protease complex shows inhibition by deformation. Nature **407**: 923-926

Karimi M, Inzé D, Depicker A (2002) GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci **7**: 193-195

Kerppola TK. (2010) Biomolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. Annu Rev Biophys. **37**: 465–487.

Lampl N, Alkan N, Davydov O, Fluhr R (2013) Set-point control of RD21 protease activity by AtSerpin1 controls cell death in Arabidopsis. Plant J **74**: 498-510

Lampl N, Budai-Hadrian O, Davydov O, Joss TV, Harrop SJ, Curmi PMG, Roberts TH, Fluhr R (2010) *Arabidopsis* AtSerpin1, crystal structure and *in vivo* interaction with its target protease RESPONSIVE TO DESICCATION-21 (RD21). J Biol Chem **285**: 13550-13560

Lease KA, Walker JC (2006) The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol **142**: 831-838

Liu J-X, Srivastava R, Che P, Howell SH (2007) Salt stress responses in Arabidopsis utilize a signal transduction pathway related to endoplasmic reticulum stress signaling. Plant J **51**: 897-909

Liu J-X, Srivastava R, Howell S (2009) Overexpression of an Arabidopsis gene encoding a subtilase (AtSBT5.4) produces a clavata-like phenotype. Planta **230**: 687-697

Matsubayashi Y (2011) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol **52**: 5-13

Matsuzaki Y, Ogawa-Ohnishi M, Mori A, Matsubayashi Y (2010) Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. Science **329**: 1065-1067

Meng L, Buchanan BB, Feldman LJ, Luan S (2012) CLE-like (CLEL) peptides control the pattern of root growth and lateral root development in *Arabidopsis*. Proc Natl Acad Sci USA **109**: 1760-1765

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in *Arabidopsis* development: how cells communicate over a short distance. Plant Cell **24**: 3198-3217

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis. Plant J **55**: 152-160

Rautengarten C, Steinhauser D, Büssis D, Stintzi A, Schaller A, Kopka J, Altmann T (2005) Inferring hypotheses on functional relationships of genes: Analysis of the *Arabidopsis thaliana* subtilase gene family. PLoS Comput Biol **1**: e40

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol **17**: 1030-1032

Schaller A, Stintzi A, Graff L (2012) Subtilases - versatile tools for protein turnover, plant development, and interactions with the environment. Physiol Plant **145**: 52-66

Srivastava R, Liu J-X, Guo H, Yin Y, Howell SH (2009) Regulation and processing of a plant peptide hormone, AtRALF23, in Arabidopsis. Plant J **59**: 930-939

Srivastava R, Liu J-X, Howell SH (2008) Proteolytic processing of a precursor protein for a growth-promoting peptide by a subtilisin serine protease in Arabidopsis. Plant J **56**: 219-227

Stahl Y, Simon R (2012) Peptides and receptors controlling root development. Phil Trans R Soc B **367**: 1453-1460

Stührwohldt N, Dahlke RI, Steffens B, Johnson A, Sauter M (2011) Phytosulfokine-α controls hypocotyl length and cell expansion in *Arabidopsis thaliana* through phytosulfokine receptor 1. PLoS ONE **6**: e21054

Tamaki T, Betsuyaku S, Fujiwara M, Fukao Y, Fukuda H, Sawa S (2013) SUPPRESSOR OF LLP1 1-mediated C-terminal processing is critical for CLE19 peptide activity. Plant J **76**: 970-81

Tsiatsiani L, Gevaert K, Van Breusegem F (2012) Natural substrates of plant proteases: how can protease degradomics extend our knowledge? Physiol Plant **145**: 28-40

Turk B (2006) Targeting proteases: successes, failures and future prospects. Nat Rev Drug Discov **5**: 785-799

Van Aken O, Pečenková T, van de Cotte B, De Rycke R, Eeckhout D, Fromm H, De Jaeger G, Witters E, Beemster GTS, Inzé D, Van Breusegem F (2007) Mitochondrial type-I prohibitins of *Arabidopsis thaliana* are required for supporting proficient meristem development. Plant J **52**: 850-864

Van Leene J, Hollunder J, Eeckhout D, Persiau G, Van De Slijke E, Stals H, Van Isterdael G, Verkest A, Neirynck S, Buffel Y, De Bodt S, Maere S, Laukens K, Pharazyn A, Ferreira PCG, Eloy N, Renne C, Meyer C, Faure J-D, Steinbrenner J, Beynon J, Larkin JC, Van de Peer Y, Hilson P, Kuiper M, De Veylder L, Van Onckelen H, Inzé D, Witters E, De Jaeger G (2010) Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. Mol Syst Biol **6**: 397

Van Leene J, Stals H, Eeckhout D, Persiau G, Van De Slijke E, Van Isterdael G, De Clercq A, Bonnet E, Laukens K, Remmerie N, Henderickx K, De Vijlder T, Abdelkrim A, Pharazyn A, Van Onckelen H, Inzé D, Witters E, De Jaeger G (2007) A tandem affinity purification-based technology platform to study the cell cycle interactome in *Arabidopsis thaliana*. Mol Cell Proteomics **6**: 1226-1238

Vercammen D, Belenghi B, van de Cotte B, Beunens T, Gavigan J-A, De Rycke R, Brackenier A, Inzé D, Harris JL, Van Breusegem F (2006) Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. J Mol Biol **364**: 625-636

Whitford R, Fernandez A, De Groodt R, Ortega E, Hilson P (2008) Plant CLE peptides from two distinct functional classes synergistically induce division of vascular cells. Proc Natl Acad Sci USA **105**: 18625-18630

Whitford R, Fernandez A, Tejos R, Cuéllar Pérez A, Kleine-Vehn J, Vanneste S, Drozdzecki A, Leitner J, Abas L, Aerts M, Hoogewijs K, Baster P, De Groodt R, Lin Y-C, Storme V, Van de Peer Y, Beeckman T, Madder A, Devreese B, Luschnig C, Hilson P (2012) GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses. Dev Cell **22**: 678-685

SUPPLEMENTAL INFORMATION

Table S1. Subtilase (*sbt*) mutant genotypes of Arabidopsis

Gene name	AGI code	Mutant ID	FST position	Homozygous	Heterozygous	Suppression of agravitropic root phenotype
AtSBT1.1	At1g01900	SALK_033704	Exon	+		No
		SALK_017912	Exon	+		No
AtSBT1.2	At1g04110	SALK_035559	300-UTR-5'	+		No
AtSBT1.3	At5g51750	SALK_011867	Exon		+	No
AtSBT1.4	At3g14067	SALK_054778	Exon	+		No
		SALK_063823	Exon	+		No
AtSBT1.5	At3g14240	SALK_032651	Exon	+		No
AtSBT1.6	At4g34980	GK_270F06	Exon	+		No
AtSBT1.7	At5g67360	GK_140B02	Exon	+		No
AtSBT1.8	At2g05920	GK_168E04	Exon	+		No
		SALK_020799	Exon	+		No
AtSBT1.9	At5g67090	SALK_009925	Exon		+	No
		SALK_009917	Exon	+		No
		GK_100G11	300-UTR-3'	+		No
AtSBT2.1	At1g30600	GK_202H08	Exon	+		No
		SALK_091134	Exon	+		No
AtSBT2.2	At4g20430	SALK_150020	Exon	+		No
		SALK_013152	Exon	+		No
AtSBT2.3	At5g44530	GK_081A06	Intron	+		No
		SALK_022324	Exon	+		No
AtSBT2.5	At2g19170	FLAG_181B06	Exon	+		No
AtSBT2.6	At4g30020	GK_125A08	Exon	+		No
		SALK_068944	Exon	+		No
AtSBT3.1	At4g21323	GK_069E04	Exon	+		No
AtSBT3.2	At1g32970	SALK_001743	1000 p	+		No
AtSBT3.3	At1g32960	SALK_086092	Exon	+		No
		SALK_107460	Exon	+		No
AtSBT3.4	At1g32950	SALK_040245	Exon	+		No
		SALK_058032	Intron	+		No
AtSBT3.5	At1g32940	GK_672C08	Intron	+		No
		SAIL_400_F09	Intron	+		No
AtSBT3.6	At4g10550	SALK_104806	Intron	+		Mild
AtSBT3.7	At4g10510	SALK_081645	Exon	+		No
AtSBT3.8	At4g10540	GK_226F04	Exon	+		No

AtSBT3.9	At4g10520	SALK_048279	Intron	+	No
AtSBT3.10	At4g10530	SALK_014429	Intron	+	No
AtSBT3.11	At5g11940	SALK_067858	Exon	+	No
AtSBT3.12	At4g21326	SALK_037231	Exon	+	Mild
AtSBT3.13	At4g21650	SALK_082160	Exon	+	No
AtSBT3.14	At4g21630	SALK_127987	Exon	+	No
AtSBT3.15	At4g21640	SALK_064593	Exon	+	No
AtSBT3.16	At1g66210	SALK_009433	Exon	+	Mild
		SALK_004741	Exon	+	Mild
AtSBT3.17	At1g66220	SALK_040473	Exon	+	No
		SALK_070765	Exon	+	No
AtSBT3.18	At4g26330	GK_360C11	Exon	+	No
AtSBT4.1	At2g39850	SALK_016756	Intron	+	No
		SALK_038521	Exon	+	No
AtSBT4.2	At4g15040	SALK_024853	Exon	+	No
AtSBT4.3	At5g59190	SALK_149055	Exon	+	No
		SALK_075909	Exon	+	No
AtSBT4.4	At5g59100	SALK_016547	1000 p	+	No
AtSBT4.5	At3g46840	GK_251C03	Exon	+	No
		SALK_078286	Intron	+	No
AtSBT4.6	At3g46850	SALK_091683	1000 p	+	No
AtSBT4.7	At5g58820	SALK_013603	Exon	+	No
AtSBT4.8	At5g58830	GT_5_112073	Exon	+	No
AtSBT4.9	At5g58840	SALK_060155	Exon	+	No
AtSBT4.10	At5g58810	SALK_119237	Exon	+	No
AtSBT4.11	At5g59130	GK_132H10	Exon	+	No
AtSBT4.12	At5g59090	GK_239B04	Exon	+	No
AtSBT4.13	At5g59120	SALK_009191	Intron	+	No
AtSBT4.14	At4g00230	SALK_019254	Exon	+	No
AtSBT4.15	At5g03620	SALK_063258	Exon	+	No
AtSBT5.1	At1g20150	SALK_017993	Exon	+	No
AtSBT5.2	At1g20160	SALK_012112	Exon	+	No
		SALK_012113	Exon	+	No
AtSBT5.3	At2g04160	SALK_051293	Intron	+	No
AtSBT5.4	At5g59810	GK_099F02	Exon	+	No
AtSBT5.5	At5g45640	SALK_107233	Exon	+	No
AtSBT5.6	At5g45650	GK_074B04	Intron	+	No
AtSBT6.1	**At5g19660**	**SALK_111474**	Exon	+	**Strong**
		SALK_020530	Exon	+	**Strong**
AtSBT6.2	**At4g20850**	**SALK_085776**	Exon	+	**Strong**

FST, flanking sequence tag; GK, GABI-Kat (http://www.gabi-kat.de/); SALK, http://signal.salk.edu/cgi-bin/homozygotes.cgi; UTR, untranslated region; 1000p: T-DNA insertion position within 1,000 bp upstream of promoter.

Table S2. *GLV1* transcript fold induction in transformed *sbt6* T-DNA mutant lines

Genotype	Transcript ratio relative to WT
WT	1.0
GLV1^{OE}	142.8
GLV1^{OE} *sbt6.1-1* line 1[a]	192.4
GLV1^{OE} *sbt6.1-1* line 2	4.7
GLV1^{OE} *sbt6.1-1* line 3	20.8
GLV1^{OE} *sbt6.1-1* line 4	11.6
GLV1^{OE} *sbt6.1-1* line 5	22.9
GLV1^{OE} *sbt6.1-1* line 6	95.2
GLV1^{OE} *sbt6.1-1* line 7	24.7
GLV1^{OE} *sbt6.1-1* line 8	2.6
GLV1^{OE} *sbt6.2* line 1[a]	39.3
GLV1^{OE} *sbt6.2* line 2	17.6

[a] Lines selected for further analysis.

Table S3. Propeptides and observed proteolytic products according to m/z for singly charged ions

Propeptide	Synthetic sequences*	Calculated *m/z*	Detected *m/z*	Cleaved peptides	Calculated *m/z*	Detected m/z
RALF23	SEIN**RRIL**ATRRYI	1761.0242	1761.4808	SEINRRIL	1000.5898	1000.8165
				ATRRYI	779.4522	779.6323
GLV1-1	NGGE**RRRAL**GGVE	1370.7247	1371.0605	NGGERRRAL	1028.5708	1028,7959
				NGGERRRA	915.4867	915.6780
GLV1-2	CASAA**RRLR**SHKHH	1629.8615	1630.7068	CASAARRLR	1003.5578	1003.9682
GLV1-3	ASAA**RRLR**SHKH	1389.7934	1390.0014	ASAARRLR	900.5486	900.6270

* Amino acids underlined correspond to reported subtilase canonical cleavage sites.

Table S4. Proteins identified after TAP purification with NTAP-Serpin1 expressed in Arabidopsis cell suspension culture

Characteristics[a]	AT1G47710	AT5G19660
Description	Serpin1	SBT6.1
# Found/three experiments	03/mrt	02/mrt
Molecular mass (kDa)	42.7	116.7
Peptide count	16	17
Protein coverage (%)	41	16
Protein score	491	163
Protein score expectation value	2.60E-45	1.60E-12
Best Ions score	108	43
Best Ions score expectation value	2.20E-10	1.30E-03

Peptide mass spectrometry was performed with the 4800 MALDI TOF/TOF™ Proteomics analyzer (AB SCIEX) and matching proteins were identified with the search engine Mascot version 2.1 (Matrix Science) using the TAIR8 database. Known TAP background proteins were filtered out.

[a] Peptide count, number of peptides with unique sequences matching the selected protein; Protein coverage %, percentage of protein sequence covered by assigned peptide matches; Protein score, score calculated by the Mascot search engine for each protein and based on the probability that peptide mass matches are nonrandom events, but if equal to or greater than the Mascot® Significance Level calculated for the database search, the protein match is considered to be statistically nonrandom at the 95% confidence interval. Protein score, -10*Log(*P*), where *P* is the probability that the observed match is a random event. Best Ions score, highest individual Ions Score for a given protein identification that is calculated by the Mascot search engine for each peptide matched from MS/MS peak lists and based on the probability that ion fragmentation matches are nonrandom events, but it is equal to or greater than the Mascot® Significance Level calculated for the database search, the peptide match is considered to be statistically nonrandom at the 95% confidence interval. Ions score, 10*Log(*P*), where *P* is the probability that the observed match is a random event.

Table S5. All GLV proteins except GLV5 comprise typical subtilaes target sequences

Protein	Amino acid sequence
GLV1	MSCSLRSGLVIVFCFILLLLSSNVGCASAARRLRSHKHHHHKVASLDVFNGGERRRALGGVETGEEVVVMDYPQPHRKPPIHNEKS
GLV2	MAIRVSHKSFLVALLLILFISSPTQARSLREVVRNRTLLVVEKSQESRKIRHEGGGSDVDGLMDMDYNSANKKRPIHNR
GLV3	MMRFTIIVIAFLLIIQSLEEEHILVYAHEGGEAGHKSLDYQGDQDSSTLHPKELFDAPRKVRFGRTTRAEKEQVTAMNNDSWSFKISGE HKQTNILADHDTTKNTFCKKMMIIVNDLTSLPTLEPSTSTNDMEKLARLLRDDYPIYSKPRRKPPVNNRAPDKF
GLV4	MEMKKWSYANLITLALLFLFFIILLLAFQGGSRDDDHQHVHVAIRTKDISMGRKLKSLKPINPTKKNGFEYPDQGSHDVQEREVYVELR DYGQRKYKPPVHN
GLV6	MKLIRVTLFLCALAILLLVTPTSSLQLKHPYSSPSQGLSKKIVTKMATRKLMIISSEYVMTSTSHEGSSEQLRVTSSGKSKDEEKKLSE EEEEKKALAKYLSMDYRTFRRRRPVHNKALPLDP
GLV7	MTTLSKILCVLIILLLCFSFRYSLHEDGNQQSSRDFVSTAKAIKYGDVMKKMIRGRKLMMASGEKEEAETKMKRGNRETERNSSKSVEE DGLVAYTADYWRAKHHPPKNN
GLV8	MKKWSYAKLMTSALLLVFLSIILLAFHGGSRGDNHLYDHVAIGTKDILMGRKLKDLKPKTESLKMINPKKKNGFEYSDQVSSDLSRQEV FVDMMARDYQGPKPRSKPLKNN
GLV9	MKKTSLKLMTLVLGFCFVIYLLQGPRGGSRNGDLLIARKLISLEPIETKNAARSLKDSISTDLEEEVDRLMEHEYPSPVKPRKRTPVHN GVRNRH
GLV10	MSSIHVASMILLLFLFLHHSDSRHLDNVHITASRFSLVKDQNVVSSSTSKEPVKVSRFVPGPLKHHHRRPPLLFADYPKPSTRPPRHN
GLV11	MVSIRVICYLLVFSVLQVHAKVSNANFNSQAPQMKNSEGLGASNGTQIAKKHAEDVIENRKTLKHVNVKVEANEKNGLEIESKEMVKKR KNKKRLTKTESLTADYSNPGHHPPRHN

Typical subtilase target sequences are presented in red.

Table S6. Primers used to confirm overexpression and knockout lines

Primer name	**Sequence 5′–3′**
Sbt6.1 LP	AGCGTACGAATTGGACAAATG
Sbt6.1 RP	AGGACCTGAAAGCTTAGCAGC
Sbt6.2 LP	ATTGAGGAACTGAGCAAATGG
Sbt6.1 RP	AGAAGTCTGCTAGTTTCCCGC
LBb1.3	ATTTTGCCGATTTCGGAAC
GLV1 F	ATGTATGTTGAATGTAAAAT
GLV1 R	AGACTTCTCGTTGTGGATCG
35S F	CCACTATCCTTCGCAAGACCCTTCC
qPCR-Sbt6.1 F	CCACCCCCGGGCAAGCATTT
qPCR-Sbt6.1 R	TGCAGGGTGCCATGTTGGTGG
qPCR-Sbt6.2 F	TCAAGCCGGGGGCCAACATC
qPCR-Sbt6.2 R	CGCAATTGCCCCACAGGCAGA
pSerpin1 F	ggggacaagtttgtacaaaaaagcaggctTAGTTAGTGTCATAATATCAAATG
pSerpin1 R	ggggaccactttgtacaagaaagctgggtTTTCGCCGGAGGTTGTGGTG

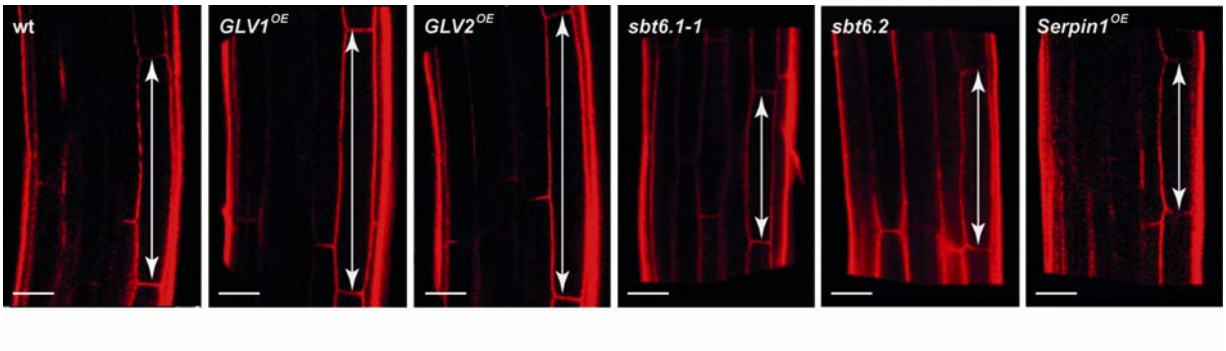


Fig. S1. Hypocotyl elongation phenotypes.

Comparison of hypocotyl epidermal cell lengths. For all genotypes, the imaged cells were the most elongated in the light-grown seedlings (5 dag). [Scale bars, 10 µm].

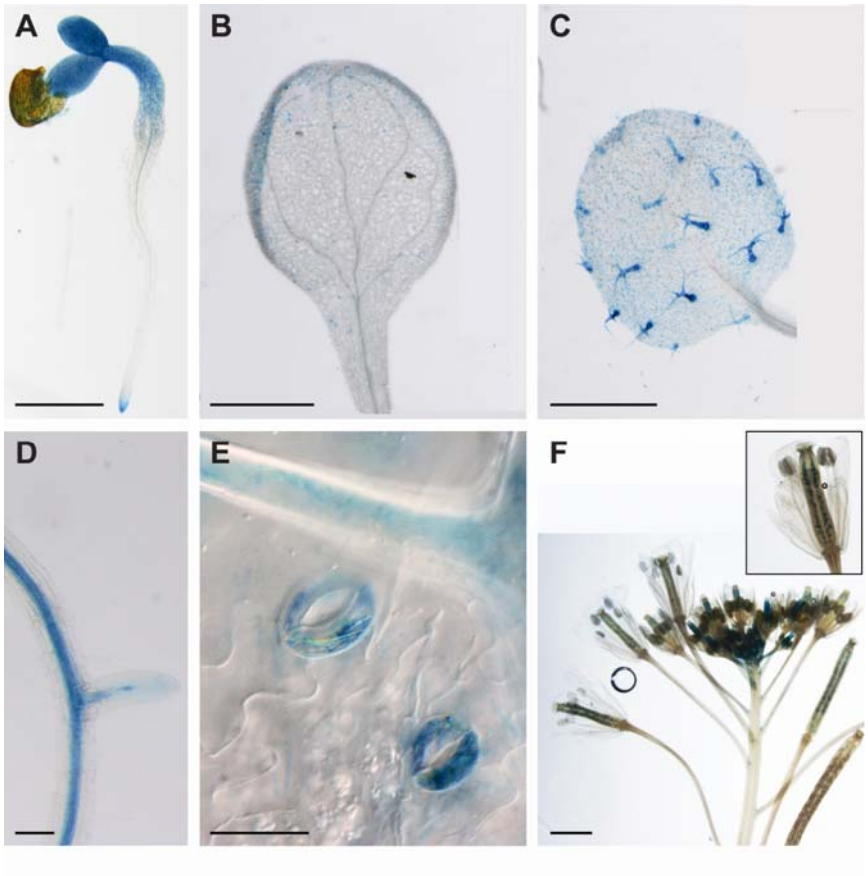


Figure S2. Transcriptional activity of pSerpin1:GUS.

A, Young seedling (3 dag). B-F, Cotyledon and first leaves, trichomes, stomata, lateral root, and inflorescence (10 dag), respectively. Bars = 1 mm (A-D and F), 25 µm (E).

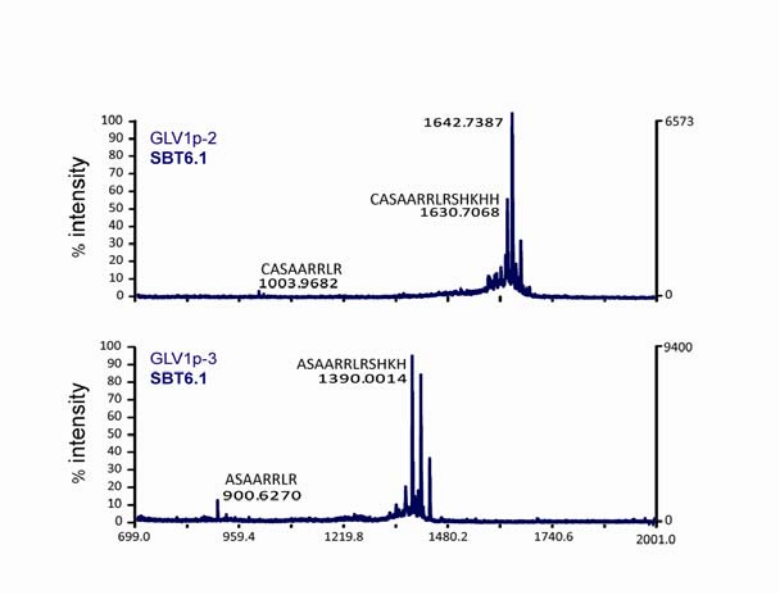


Figure S3. MALDI-TOF spectra for synthetic peptides.

Synthetic GLV1-2 and GLV1-3 peptides (Table S3) and fragments after 1 h of incubation with SBT6.1.

Chapter 4

GLV genes contribute to root development

This chapter is part of an ongoing study.

Authors' contribution: S.G. designed the experiments with the help of A.F., P.H. and T.B; S.G. performed all the experiments; S.G. wrote the manuscript with help of T.B.

Sarieh Ghorbani,[1,2] Ana Fernandez,[1,2] Pierre Hilson[1,2,3,4] and Tom Beeckman,[1,2*]

[1]Department of Plant Systems Biology, VIB, B-9052 Ghent, Belgium

[2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium

[3]INRA, UMR1318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

[4]AgroParisTech, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

[*]Corresponding author

Running head: *GLV10* peptide in root development

Abstract

Small signaling peptides are involved in numerous aspects of plant growth and development. Peptides are particularly suited for short-distance communications and their vital role in root development has been demonstrated. Members of the *GOLVEN/ROOT GROWTH FACTOR/CLE-LIKE* (*GLV/RGF/CLEL*) signaling peptide family are expressed differentially through the root system and contribute to root apical meristem maintenance and lateral root development. Morphological analysis revealed that overexpression of *GLV* genes can perturb normal growth and emergence of lateral root primordia. In the majority of gain-of-function lines, the shape of lateral root primordia is aberrant and the regular cell division pattern is disturbed. Loss-of-function of some of the *GLV* genes resulted in an increased number of lateral roots. Because of its expression during the early stages of lateral root development the potential function of *GLV10* was analyzed in more detail. As is the case for most of the *GLV* genes, primordia in *GLV10* gain-of-function mutants lost the organized cell division pattern and produced aberrant primordia. This deviation from the normal lateral root development process resulted in a dramatic reduction in the number of emerged lateral roots that, in contrast, was increased in partial loss-of-function *GLV10* mutant lines. Furthermore, *GLV10* was expressed in the root apical meristem and loss-of-function *GLV10* mutant lines caused a decrease in the root apical meristem length. In summary, our results indicate that different *GLV* genes contribute to

lateral root organogenesis and that *GLV10* regulates lateral root primordia development. Additionally, together with *GLV5*, *GLV7,* and *GLV11*, *GLV10* might be involved in root apical meristem maintenance.

Introduction

Plants utilize signaling peptides in cell-to-cell communications to orchestrate growth and development (Matsubayashi, 2011a; Murphy et al., 2012). Over the past decades, several small signaling peptide families have been detected in plants that coordinate diverse aspects of plant growth and development (Fletcher, 1999; Yang et al., 1999; Yang et al., 2000; Pearce et al., 2001b; Casson et al., 2002; Butenko et al., 2003; Hara et al., 2007; Ohyama et al., 2008; Matsuzaki et al., 2010; Whitford et al., 2012). The *GOLVEN (GLV)/ ROOT GROWTH FACTOR (RGF)/CLE-LIKE (CLEL)* family has been identified coincidentally through three independent *in silico* studies, hence the different nomenclature (Matsuzaki et al., 2010; Meng et al., 2012; Whitford et al., 2012). For clarity, these peptides will be referred to according to their *GLV* nomenclature.

The genome of the model species *Arabidopsis thaliana* encodes 11 *GLV* genes. Further studies revealed their potential role in a number of developmental programs (Fernandez et al., 2013b), one of which is root apical meristem (RAM) maintenance. *GLV11* acts posttranslationally on the *PLETHORA (PLT)* transcription factors that are involved in stem cell maintenance to define their expression levels and patterns (Matsuzaki et al., 2010). Abnormal expression of *GLV* genes hampers the formation of the auxin gradient in gravistimulated seedlings. The *GLV* signaling has been proposed to modulate the auxin gradient by regulating the trafficking dynamics of the auxin efflux carrier PIN-FORMED2 (PIN2) (Whitford et al., 2012). Ectopic overexpression of *GLV* genes as well as supplementing growth media with mature GLV peptides resulted in impaired root and shoot gravitropic responses (Whitford et al., 2012). Moreover, *GLV* genes might be involved in lateral root (LR) and root hair development (Meng et al., 2012; Fernandez et al., 2013a).

The successive stages of LR development have been described exhaustively (Malamy and Benfey, 1997). LRs arise from pericycle cells adjacent to xylem poles in *Arabidopsis* (Dolan et al., 1993; Charlton, 1996; Osmont et al., 2007). LR initiation is recognizable by migration of the nuclei of two founder cells towards the common cell wall. Soon after, the first asymmetric cell divisions take place that yield daughter cells with different sizes. Afterwards, several rounds of anticlinal divisions start forming a single layer consisting of up to 10 cells (Malamy and Benfey, 1997; Dubrovsky et al., 2001; De Rybel et al., 2010). During development of lateral root

primordia (LRP), these cells divide periclinally to form a primordium with an outer and inner layer. Later, cells undergo several precise anticlinal and periclinal divisions to develop a dome-shaped primordium. Eventually, the newly formed LR emerges through the main root (Malamy and Benfey, 1997).

Our understanding of the molecular mechanisms that control LR initiation is steadily improving (Benkova and Bielach, 2010). However, the discovery that complex regulatory networks involving secreted peptide hormones also take part in these processes is recent. The systematic study of the *GLV* expression in *Arabidopsis* revealed that 10 out of 11 *Arabidopsis GLV* genes are transcribed in lateral roots, in restricted domains, and at distinct stages of development (Fernandez et al., 2013a). The specificity of their expression patterns suggests that *GLV* genes may control LR development in distinct ways (Fernandez et al., 2013a). Furthermore, the diversity of the *GLV* expression patterns and the range of LR emergence inhibition resulting from the overexpression of different *GLV* genes (Fernandez et al., manuscript in preparation; this study) imply that different members of the family control distinct processes during LR initiation. Further studies are needed to investigate their precise role in this process.

To better understand the role of *GLV* genes in root development, we investigated the morphological changes in LRP upon ectopic overexpression of *GLV* genes. Furthermore, the role of the *GLV10* gene in primary root and LR development was studied. Genetic and morphological analyses showed that *GLV10* is active in LR development and root apical meristem functioning.

Results

GLV gain-of-function affects lateral root development

Recently, several *GLV* genes have been suggested to play a role in LR development purely on the basis of their expression patterns (Fernandez et al., 2013a). Here, we investigated the root phenotypes of *GLV* gain-of-function lines (GLV^{OE}) in an attempt to verify their proposed role in root development. Strongly reduced lateral root densities, determined by the number of outgrown lateral roots per unit root length, were recorded in two analyzed independent lines of $GLV2^{OE}$, $GLV5^{OE}$, $GLV6^{OE}$, $GLV7^{OE}$, $GLV10^{OE}$, and $GLV11^{OE}$, with a complete absence of emerged lateral roots in the $GLV6^{OE}$ line (Figure 1A). Closer examination of these roots at the microscopical level revealed the presence of multiple sites of cell proliferation along the pericycle. Division patterns clearly deviated from the normal division pattern associated with primordia development in the wild type (WT), which is characterized by the formation of a dome-shaped structure from stage IV onwards. On the contrary, in $GLV2^{OE}$, $GLV5^{OE}$, $GLV6^{OE}$, $GLV7^{OE}$, $GLV10^{OE}$, and $GLV11^{OE}$ roots, some primordia underwent extra anticlinal cell divisions, thereby expanding the division site that became elongated and incapable of forming dome-shaped structures (Figure 1). Such misshapen primordia seemed to be arrested between stage II and V. Besides the arrested phenotype, some merged or closely positioned primordia were observed in $GLV10^{OE}$ and $GLV11^{OE}$ lines that had the most dramatic phenotype in this respect. However, no quantification has been done yet.

In addition to *GLV5*, *GLV6*, *GLV7*, *GLV10,* and *GLV11*, which are expressed at early stages of LRP development, $GLV2^{OE}$ also showed defects at the early stages of LR growth and development (Figure 1A and C) although its expression normally occurrs only after the establishment of the LR meristem and in the boundary cells of LRP (Figure 1B). Most of the LRP in $GLV2^{OE}$ failed to form a dome-shape and primordia adjacent to each other were also visible (Figure 1C).

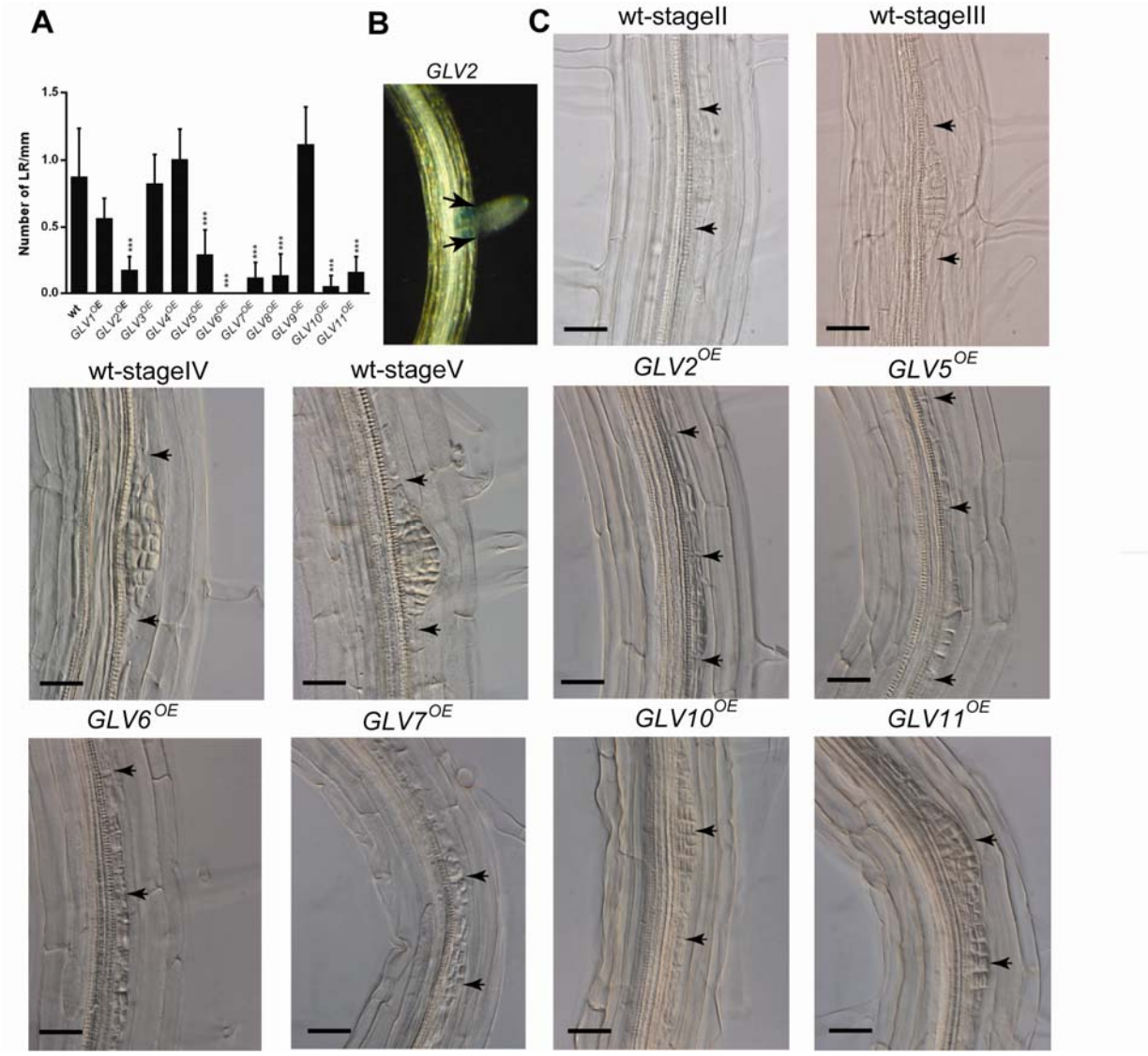


Figure 1: ***GLV*** **gain-of-function effects on lateral root development**

A, Quantification of the emerged LR density in GLV^{OE} lines. For simplicity, results for one of the two analyzed independent lines is shown. Error bars show 95% confidence interval (n = 20-35). B, *GLV2*-driven GUS activity in boundary cells of LRs, imaged with differential interference contrast (DIC) microscopy. Arrowheads point to the *GLV2-GUS* signal in boundary cells of LRs. C, DIC images of GLV^{OE} genes and wild type (WT) primordia at stage II, III, IV, and V. WT LRP form a dome-shaped structure from stage IV onwards. *GLV2, GLV5, GLV6, GLV7, GLV10,* and *GLV11* overexpression resulted in the formation of multiple cell proliferation sites along the pericycle and abnormal primordia. Some primordia underwent extra anticlinal cell divisions that formed large and stretched dividing cells and failed to form a dome-shaped structure. Arrows in the WT stage II to stage V indicate borders of normal and aberrant primordia in *GLV*-overexpressing mutant lines. Scale bars represent 20 µm.

GLV loss-of-function promotes lateral root development

The carboxyl-termini of the different GLV precursors contain the bioactive secreted peptides and show high sequence similarities (Fernandez et al., 2013b), suggesting that the *GLV* genes might be functionally redundant and, hence, that gain-of-function *GLV* plants often exhibit similar phenotypes (Fernandez et al., 2013b). To get insight into the contribution of each *GLV* to the development of the primary root and LRs, we studied all available *GLV* loss-of-function mutants and generated RNA interference knock-down lines in which the normal function of the *GLV* genes is perturbed.

Loss-of-function of several *GLV* genes increased the number of non-emerged LRs, However, knock-down seedlings of *GLV3* (*amiRglv3*) showed increased numbers of both emerged and non-emerged LRs (Figure 2) and a significant increase in the number of stage-I primordia. More stage-I primordia were also found in *amiRglv4*, *glv7* and *glv8* mutants, whereas the number of stage-II primordia was higher than that of the control wild-type plants in *amiRglv1* and *amiRglv3* mutants. The number of emerged (E) LRs in *amiRglv3* lines and the number of non-emerged (NE) LRs in *amiRglv1*, *amiRglv3*, *amiRglv4* and *glv7* was higher than that of control lines. Finally, *amiRglv1*, *amiRglv3*, *amiRglv4*, *glv7* and the *glv5glv7* double mutant differed significantly in the total (E+NE) number of LRP when compared to control wild-type plants (Figure 2 and Figure S1). In summary, our data show that an increase in the number of LRs generally occurs either at stage I, stage II and/or at the emergence stage in *glv* loss-of-function or knock-down mutants. No changes could be detected in the number of LRP at stages III to VII.

GLV10 gain-of-function results in short cells along the pericycle

Given its early expression during LR formation (from stage II onwards as reported by Fernandez et al., 2013a) and the dramatic effect of overexpression on this process (Figure 1A and C), we decided to characterize in more detail the *GLV10* gain-of-function and knock-down root phenotype. First, we reinvestigated the *GLV10* expression pattern accurately during different stages of LR formation. The *GLV10* transcriptional signal was first observed at stage II. After the first appearance of the *GLV10* transcriptional signals, transcription occurred until the establishment of the LR meristem and later on in the emerged LRs (Figure 3A).

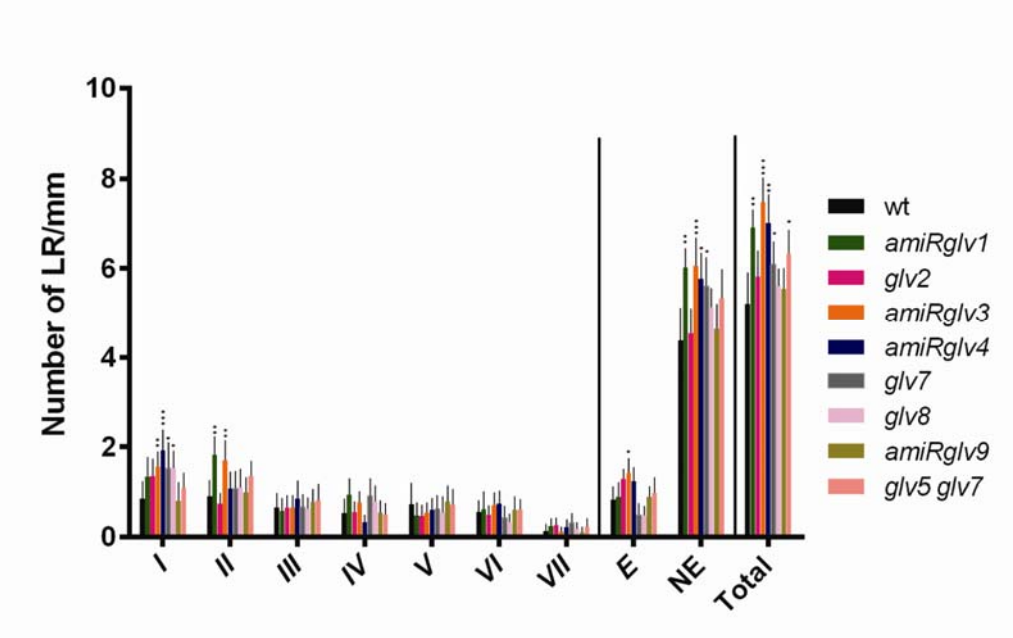


Figure 2: Distribution of LR developmental stages in control and *GLV* loss-of-function and knock-down lines.

Error bars show the 95% confidence interval. For simplicity, results for one of two analyzed independent mutant lines is presented (see Materials and Methods). Asterisks indicate significant differences compared to the wild type (P<0.001). I to VII, primordium stages; NE, nonemerged primordia; E, emerged LRs; total, total number of LRs.

The *GLV10* transcriptional signal appears first in just two or three cells in the LRP centers, but in the mature LRs, it was observed in the newly established RAM, in the QC, and columella cells. This expression was similar to the one reported for its expression in the main root (Fernandez et al., 2013a). In addition, *GLV10* was expressed in the LR boundary cells (Figure 3A).

To understand how overproduction of *GLV10* abolishes normal LR growth, we studied possible morphological changes in LRP cells in *GLV10* gain-of-function lines. Close microscopical analysis revealed that overproduction of *GLV10* generated very short protoxylem pole pericycle cells along the root compared to those of wild-type plants (Figure 3B and C). Whereas the majority of the wild-type pericycle cells were longer than 60 µm, $GLV10^{OE}$ pericycle cells were on average shorter than 29 µm (Figure 3B). Nevertheless, *GLV10* is not expressed during LRI early stages and the short pericycle cell phenotype might be the result of GLV10 binding to another than its own receptor such as the GLV6 receptor. In summary,

transcriptional and gain-of-function data during LRP development suggest a potential role for *GLV10* during post-initiation developmental stages of LRs.

GLV10 loss-of-function affects primary and lateral root development

Next, we searched for corresponding loss-of-function mutants, but could not detect any T-DNA insertion mutant in the *GLV10* open reading frame (ORF) in the available T-DNA insertion mutant collections. Therefore, we generated transgenic lines carrying a transgene cassette that encoded the artificial microRNA precursor (*amiRglv10*) to silence the *GLV10* expression. Two different amiRNA constructs were designed to target the *GLV10* transcript by means of the Web MicroRNA Designer Tool (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi). We placed the *amiRglv10s* under the control of either the strong *35S CaMV* promoter or the *GLV10* endogenous promoter. As the expression pattern of *GLV10* is similar to that of *GLV5*, *GLV7, GLV11* in the RAM of *Arabidopsis* (Fernandez et al., 2013a), *amiRglv10* lines were transformed into Col-0 and *glv5,glv7,glv11* genotypes. These combinations resulted in eight different combinations (Table S1).

Between the eight different T3 homozygous genotypes, the silencing rate in the lines under control of the *GLV10* endogenous promoter was higher than those under control of the strong *35S CaMV* promoter (Figure 4A; Figure S2). In total, we could detect two T3 homozygous *amiRglv10* lines in the *glv5,7,11* background and one line in the Col-0 background with a decreased *GLV10* transcription level compared to that of the corresponding control lines (Figure 4A). Among all the tested homozygous lines, the *GLV10amiRglv10-2 glv5,7,11* line showed the strongest silencing in *GLV10* transcription level. The transcription level of *GLV10* in the *GLV10amiRglv10-2 glv5,7,11* line decreased almost 90% compared to that of the *glv5,glv7,glv11* triple mutant, but no strong decrease could be detected in the wild-type background. Nevertheless, to avoid that lines with strong phenotypes in all generated lines could have been omitted, we analyzed the LR density and root length in 45 independent T3 homozygous lines in the Col-0 background and 73 independent T3 homozygous lines in the *glv5,7,11* background both under control of the strong *35S CaMV* and endogenous *GLV10* promoters (Figure S3-1 and S3-2). For further analyses, we chose to use the *GLV10amiRglv10-2 glv5,7,11* line (Figure4A), which, hereafter, will be referred to as *GLV10amiRglv10 glv5,7,11*.

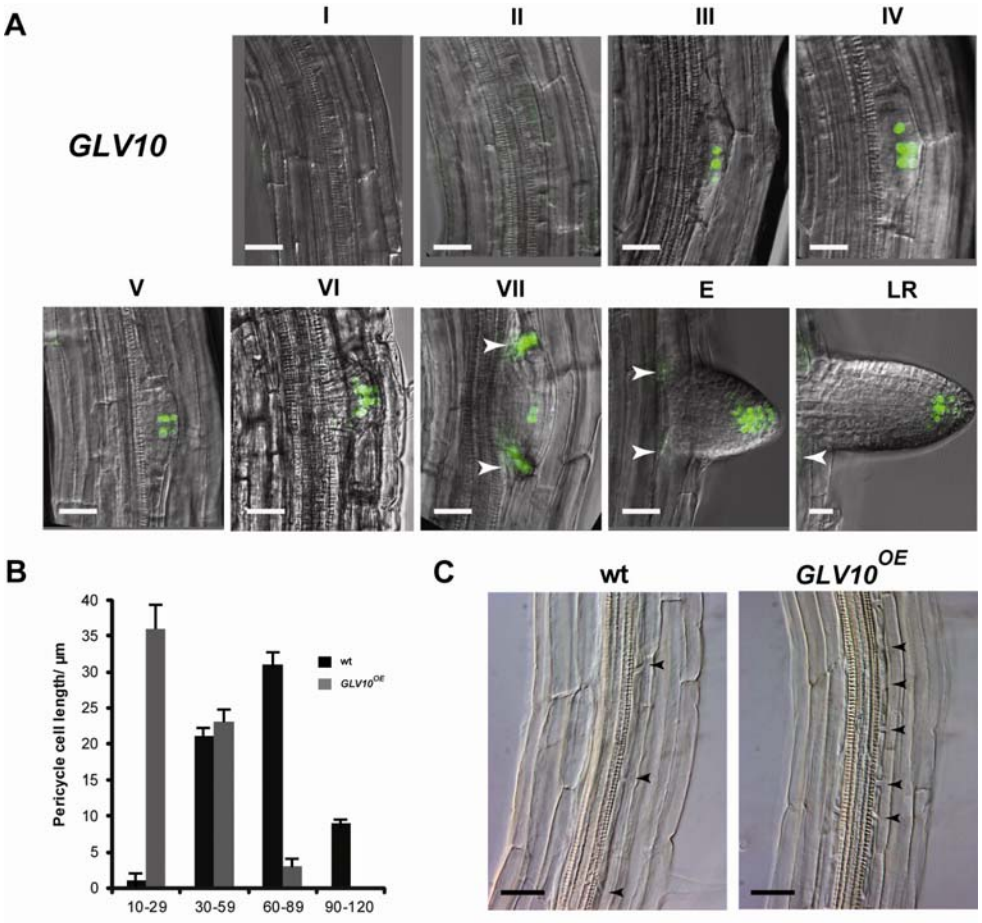


Figure 3: LR development and root growth promoted by loss-of-function of *GLV* genes

A, Expression of *GLV10* during the early stages of LR development. The images show the expression pattern of the *GLV10* gene at different developmental stages according to Malamy and Benfey (1997). Stages are indicated on top of each picture. E, emerged LR; LR, mature LR with established apical meristem. Arrowheads point the *GLV10-NLS-2XGFP* signal in boundary cells of LRs. B, Quantification of pericycle cell length (µm) in $GLV10^{OE}$ and wild-type (WT) roots (n=63). Cell length was measured for the cells that remained between the two youngest existing primordia in seedlings 10 days after germination (dag). C, DIC images of representative $GLV10^{OE}$ and wild-type pericycle cells. *GLV10* overproduction resulted in the formation of shorter cells along the pericycle pole than those in the the WT. Arrowheads point the boundary of the pericycle cells in 10-dag seedlings. Scale bars represent 20 µm.

To investigate whether *GLV10* silencing affects LR development, we studied the possible morphological changes at different stages of LRP. *GLV10:amiRglv10 glv5,7,11* primordia divided and grew normally. In T3 homozygous *GLV10:amiRglv10 glv5,7,11* plants, the number

of emerged LRs was higher than that of the *glv5,7,11* control line (Figure 4D) and there was a slight increase in the number of stage-II and stage-III LRP, but not significant compared to the control line.

As the expression of *GLV10* is similar to that of *GLV5*, *GLV7, GLV11* in the RAM of *Arabidopsis* (Fernandez et al., 2013a) and the latter are redundantly involved in the RAM maintenance (Matsuzaki et al., 2010), we speculated that *GLV10* functioned similarly in the RAM regulation. To examine this hypothesis, we analyzed the effect of *GLV10* malfunction on RAM development in knock-down mutant lines. *GLV10* silencing resulted in a shorter meristem size than that of the *glv5,glv7,glv11* triple mutant control line (Figure 4C and E). Consistent with this result, the root length of *amiRglv10 glv5,7,11* mutants lines was shorter than that of the *glv5,glv7,glv11* triple mutants (Figure 4B).

In summary, our results indicate that *GLV10* is involved in the regulation of both primary root and LR development in *Arabidopsis*. Besides *GLV5*, *GLV7,* and *GLV11*, *GLV10* is also involved in the RAM regulation of *Arabidopsis*.

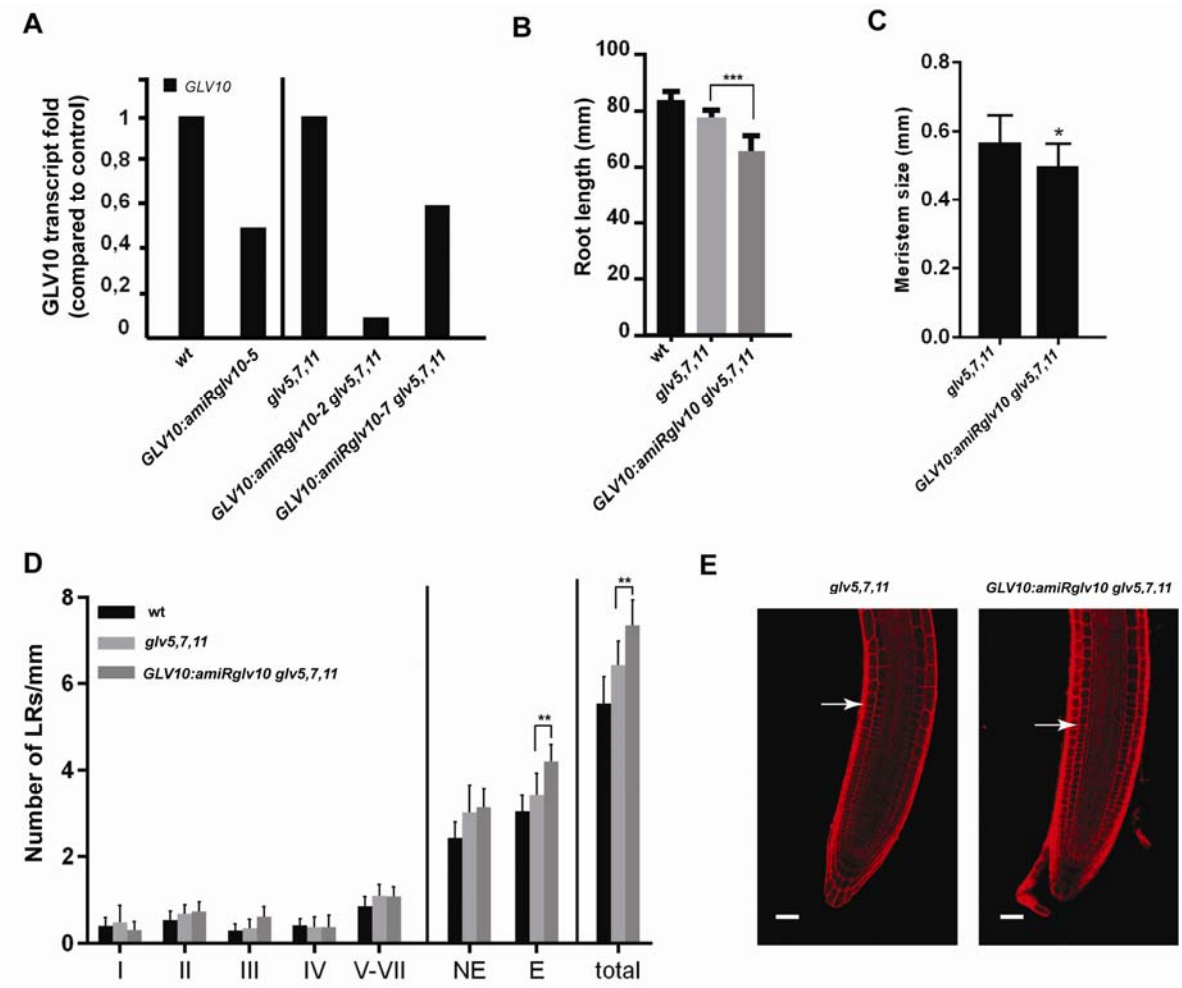


Figure 4: Root and LR development promoted by *GLV10* loss-of-function

A, qRT-PCR analysis of overexpression of *amiRglv10* under the control of the endogenous *GLV10* promoter in Col-0 and *glv5,7,11* background compared to the wild type and *glv5,7,11,* (mean transcript fold decrease ± confidence interval [CI]), For further analyses, we chose to use the GLV10amiRglv10-2 glv5,7,11 line. B, Quantification of root length (mm) (n=20-27). Root length was measured in 12-dag seedlings. C, Quantification of root meristem size (mm) (n=12-22). RAM size was measured in 5-dag seedlings. Statistical differences between *amiRglv10 glv5,7,11* with the control line *glv5,7,11* were assessed with Student's *t*-tests. D, Distribution of LR developmental stages in 12-dag seedlings of control and *GLV* knock-down lines. I to VII, primordium stages; NE, nonemerged primordia; E, emerged LRs; total, total number of LRs. Results for one of two analyzed independent experiments are shown (see Materials and Methods). Error bars show 95% confidence interval. Asterisks indicate significant differences compared to the wild type ($P<0.001$). E, Representative RAM lengths in *glv5,7,11* and *GLV10:amiRglv10 glv5,7,11*. Scale bars represent 20 µm. Arrowheads mark borders of the meristematic zone.

Discussion

As GLV peptides have been suggested to be involved in numerous developmental programs, including root and LR development (Matsuzaki et al., 2010; Meng et al., 2012; Whitford et al., 2012; Fernandez et al., 2013a), we focused on the characterization of the GLV signaling peptide family during primary root and LR development. To study the potential roles of *GLV* genes in root and LR development, we analyzed the morphological changes in LRP upon the ectopic overexpression of *GLV* genes. Morphological analysis revealed that overexpression of *GLV* genes can perturb normal growth and emergence of the LRP. In the majority of the gain-of-function lines, the LRP shape is abolished and normal cell division is disturbed. In WT, up to 10 cells, which are generated through several anticlinal divisions, form the LRP base, but in *GLV* overexpression mutant lines, a higher number of pericycle cells undergo anticlinal divisions with an increased LRP width along the root axis as a consequence. Moreover, in *GLV* gain-of-function lines, sometimes two or three LRP are fused and cannot proceed to the next stage. Consequently, in gain-of-function lines most of the LRP fail to form a dome-shaped structure as seen in wild-type roots. Interestingly, only the gain-of-function *GLV* genes that had no or very few emerged LRs exhibited a disturbed LRP development at the early stages of LRP growth. The normal cell division was abolished in all these lines. The data show that spatiotemporal regulation of cell division in LRP has a great impact on the fate of a primordium and suggest a possible role of GLV peptides in the coordination of cell division at the early stages of LRP development. One of the possible hypotheses is that GLV peptides might act as an upstream component of some of the cell cycle genes, which, if perturbed, can deregulate the precision of cell divisions during LRP development.

Loss-of-function of some of the *GLV* genes resulted in an increased number of LRs, but this phenotype has been observed also for genes that are not expressed in LRP, namely *GLV1*. Additionally, some genes even if not expressed at a given developmental stage, still give rise to effects the frequency of some stages; for example, the *GLV4* loss-of-function has more stage-I and the *GLV3* more stage-II primordia than the wild type (Figure 2). This observation might hint at a possibly indirect effect of these genes on LRP development and *GLV1*, *GLV4* and *GLV3* might regulate other genes that are involved directly in LRP growth.

LRs initiate when pericycle cells accumulate auxin, thereby obtaining a founder cell status that will trigger asymmetric cell divisions to form a new LRP (Van Norman et al., 2013b). In gravistimulated roots, auxin accumulation is coordinated by *GLV3* via regulation of PIN2 localization (Whitford et al., 2012) and *GLV11* defines *PLT1* and *PLT2* expression in RAM (Matsuzaki et al., 2010). Recently, auxin has been demonstrated to regulate the *PLT* function through the *AUXIN RESPONSIVE FACTOR 7 (ARF7)* and *ARF19* in LRP (Hofhuis et al., 2013) and *PIN3* has been found to play an important role in LR initiation (Marhavy et al., 2013). As one of the *GLV* genes that is expressed early during LR formation, *GLV10* overexpression caused severe defects in LRP development, resulting in misshapen primordia. *GLV10* gain-of-function mutants lost the organized cell division pattern, produced aberrant primordia and, eventually, reduced the number of emerged LRs. Consistently, the number of emerged LRP increased in loss-of-function *GLV10* mutant lines. Besides aberrant LRP, pericycle cells in $GLV10^{OE}$ are much shorter than the wild type, which might be the consequence of extra divisions in these cells. As cell division is disrupted in the pericycle cells and LRP of $GLV10^{OE}$, *GLV10* might affect the auxin maxima via the regulation of the PIN3 localization and control *PLT3*, which is expressed in LRP, in accordance with *GLV3* and *GLV11*. Further study is required to trace the auxin maxima in *GLV10* gain-of-function and knock-down lines to investigate the likely *GLV10*-dependent regulation of the local auxin concentration in LRP.

As *GLV10* is expressed in the RAM, a possible function might be expected in RAM regulation as well. Our data showed that decreases in *GLV10* transcript levels in the *glv5,glv7,glv11* triple mutant background resulted in an even shorter RAM than that of both *glv5,7,11* and the wild type. This result suggests that, in addition to *GLV5, GLV7,* and *GLV11* (Matsuzaki et al., 2010), *GLV10* also contributes to the RAM maintenance. The expression pattern of *PLT1*and *PLT2* has been shown to be regulated by *GLV11* in the RAM (Matsuzaki et al., 2010), hence, *GLV10* might also regulate their function, but detailed analysis is required to study a possible *PLT-GLV10* relation and to dissect how *GLV10* regulates RAM size and LR development.

In conclusion, we have shown that different *GLV* genes contribute to LR organogenesis and our data suggest that *GLV10* regulates the LRP development. Additionally, together with *GLV5*, *GLV7* and *GLV11*, *GLV10* is involved in the RAM maintenance.

Materials and Methods

Growth conditions

Seeds were surface-sterilized and sown on half-strength Murashige and Skoog (MS) medium (Duchefa Biochemie B.V.) complemented with 1% (w/v) agarose and 1.5% (w/v) sucrose at pH 5.8, and stratified for at least 2 days at 4°C. Seedlings were germinated in illuminated growth chambers under a 16-h light/8-h dark cycle (100-µmol m^{-2} s^{-1}) at 21°C.

Plant material

Wild-type *Arabidopsis thaliana* (L.) Heyhn. accession Colombia-0 (Col-0) was used as control, unless otherwise specified. Transgenic overexpressing, artificial microRNA and reporter lines have been described previously: *GLV1*^{OE}, *GLV2*^{OE}, *GLV3*^{OE} (Whitford et al., 2012), *GLV4*^{OE} to *GLV11*^{OE} (Fernandez et al., 2013a); *GLV1pro::NLS-GFP-GFP*, *GLV2pro::NLS-GFP- GFP* and *GLV3pro::NLS-GFP-GFP* (Whitford et al., 2012); *GLV4pro::NLS-GFP-GFP* to *GLV11pro::NLS-GFP- GFP* (Fernandez et al., 2013a);*, amiRglv1* and *amiRglv3* (Whitford et al., 2012); and *amiRglv4* and *amiRglv9 (Fernandez et al., 2013a).* The mutant lines *glv8-1* (SALK_054452), *glv5-1* (= *rgf2-1* in Matsuzaki et al., 2010) (SALK_145834) and *glv7-1* (= *rgf3-1* in (Matsuzaki et al., 2010) (SALK_053439) were obtained from the European *Arabidopsis* Stock Center (NASC). *glv11* (= *rgf1-1* in Matsuzaki et al., 2010) and *glv5glv7glv11* (= *rgf123* in Matsuzaki et al., 2010) were provided by Prof. Dr. Matsubayashi. Plants homozygous for *glv5* and *glv7* were obtained through crosses.

GLV10 was silenced by expression of an artificial microRNA (amiRNA) constructed with the online tool WMD3 (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi) (Ossowski et al., 2008). amiRNAs were constructed based on two different original sequence hints, Original hint 1 (TTAATCTGCAAATAAGCGCGG) and Original hint 2 (TCTAAGAGGATTATGAACCAT) primers used are listed in Table S2. The construct, under control of either the *35S CaMV* or the *GLV10* endogenous promoter, was transformed into Col-0 and the *glv5,7,11 (=rgf123)* mutant line by the floral dip method with the *Agrobacterium tumefaciens* strain C358C1 (Clough and Bent, 1998).

Morphological analysis

Root and RAM lengths were measured by Image J software (http://rsbweb.nih.gov/ij/). The number of emerged LRs, the total LR number and detailed number of LRs presenting different developmental stages were counted in 8- to 12-dag seedlings. All the data were normalized according to the root length. Roots were cleared based on the protocol described by (Malamy and Benfey, 1997) and analyzed and imaged with an Olympus microscope (BX51) at 40x magnification.

Histochemical and microscopic analysis

For live-cell imaging, seedlings were mounted in water with or without dye. GFP fluorescence was imaged with a 488-nm laser excitation. Emission fluorescence was captured in the frame-scanning mode alternating GFP fluorescence via a 500-/550-nm band-pass emission filter. RAMs were counterstained with propidium iodide, imaged with a 543-nm filter, and 590- to 620-nm for excitation and detection, respectively.

Gene expression analysis

Total RNA from 2-week-old roots was isolated with the TRIzol reagent (Invitrogen), followed by treatment with RNase-free DNase I (Qiagen) according to the manufacturer's instructions. The cDNA was prepared with the iScript™cDNA Synthesis Kit (Bio-Rad) from 1 μg of total RNA. For quantitative RT-PCR, 1:10 dilutions of total cDNA were used (all the primers are listed in Table S1).

Root meristem size analysis

For the root meristem size, the distance was measured between the QC and the transition zone (TZ), as indicated by the position of the first elongating cortical cell (Dello Ioio et al., 2007). Five-dag roots were stained by propidium iodide and imaged for this analysis.

Statistical tests

Means of samples were compared with one-way or two-way analysis of variance (ANOVA) (GraphPad Prism; V6.00, GraphPad Software). Results were obtained by pooling data from two or three independent biological replicates unless specified otherwise. All the error bars represent 95% confidence interval.

Acknowledgement

We thank Yoshikatsu Matsubayashi for the generous gift of the *Arabidopsis glv11 (=rgf1-1)* and *glv5glv7glv11 (=rgf123)* mutant lines. This work was supported by the Integrated Project AGRON-OMICS, in the Sixth Framework Programme of the European Commission (LSHG-CT-2006-037704).

References

Benková E, Bielach A (2010) Lateral root organogenesis -- from cell to organ. Curr Opin Plant Biol **13**: 677-683

Butenko MA, Patterson SE, Grini PE, Stenvik G-E, Amundsen SS, Mandal A, Aalen RB (2003) *INFLORESCENCE DEFICIENT IN ABSCISSION* controles floral organ abscission in Arabidopsis and identifies a novel family of putative ligands in plants. Plant Cell **15**: 2296-2307

Casson SA, Chilley PM, Topping JF, Evans IM, Souter MA, Lindsey K (2002) The *POLARIS* gene of Arabidopsis encodes a predicted peptide required for correct root growth and leaf vascular patterning. Plant Cell **14**: 1705-1721

Charlton WA (1996) Lateral root initiation. In *Plant Roots: the Hidden Half*, Y Waisel, A Eshel, U Kafkafi (Eds), 2nd ed. New York, Marcel Dekker, pp 149-173

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735-743

De Rybel B, Audenaert D, Xuan W, Overvoorde P, Strader LC, Kepinski S, Hoye R, Brisbois R, Parizot B, Vanneste S, Liu X, Gilday A, Graham IA, Nguyen L, Jansen L, Njo MF, Inzé D, Bartel B, Beeckman T (2012) A role for the root cap in root branching revealed by the non-auxin probe naxillin. Nat Chem Biol **8**: 798-805

Dello Ioio R, Scaglia Linhares F, Scacchi E, Casamitjana-Martinez E, Heidstra R, Costantino P, Sabatini S (2007) Cytokinins determine *Arabidopsis* root-meristem size by controlling cell differentiation. Curr Biol **17**: 678-682

Dolan L, Janmaat K, Willemsen V, Linstead P, Poethig S, Roberts K, Scheres B (1993) Cellular organisation of the *Arabidopsis thaliana* root. Development **119**: 71-84

Dubrovsky JG, Rost TL, Colón-Carmona A, Doerner P (2001) Early primordium morphogenesis during lateral root initiation in *Arabidopsis thaliana*. Planta **214**: 30-36

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, Hilson P (2013) Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation. Plant Physiol **161**: 954-970

Fernandez A, Hilson P, Beeckman T (2013b) GOLVEN peptides as important regulatory signalling molecules of plant development. J Exp Bot **64**: 5263-5268

Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM (1999) Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science **283**: 1911-1914

Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM (1999) Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. Science **283**: 1911-1914

Hara K, Kajita R, Torii KU, Bergmann DC, Kakimoto T (2007) The secretory peptide gene *EPF1* enforces the stomatal one-cell-spacing rule. Genes Dev **21**: 1720-1725

Hofhuis H, Laskowski M, Du Y, Prasad K, Grigg S, Pinon V, Scheres B (2013) Phyllotaxis and rhizotaxis in *Arabidopsis* are modified by three PLETHORA transcription factors. Curr Biol **23**: 956-962

Malamy JE, Benfey PN (1997) Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. Development **124**: 33-44

Marhavý P, Vanstraelen M, De Rybel B, Ding Z, Bennett MJ, Beeckman T, Benková E (2013) Auxin reflux between the endodermis and pericycle promotes lateral root initiation. EMBO J **32**: 149-158

Matsubayashi Y (2011a) Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol **52**: 5-13

Matsuzaki Y, Ogawa-Ohnishi M, Mori A, Matsubayashi Y (2010) Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. Science **329**: 1065-1067

Meng L, Buchanan BB, Feldman LJ, Luan S (2012) CLE-like (CLEL) peptides control the pattern of root growth and lateral root development in *Arabidopsis*. Proc Natl Acad Sci USA **109**: 1760-1765

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in *Arabidopsis* development: how cells communicate over a short distance. Plant Cell **24**: 3198-3217

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis. Plant J **55**: 152-160

Osmont KS, Sibout R, Hardtke CS (2007) Hidden branches: developments in root system architecture. Annu Rev Plant Biol **58**: 93-113

Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J **53**: 674-690

Pearce G, Moura DS, Stratmann J, Ryan CA Jr (2001b) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proc Natl Acad Sci USA **98**: 12843-12847

Van Norman JM, Xuan W, Beeckman T, Benfey PN (2013b) To branch or not to branch: the role of pre-patterning in lateral root formation. Development **140**: 4301-4310

Whitford R, Fernandez A, Tejos R, Cuéllar Pérez A, Kleine-Vehn J, Vanneste S, Drozdzecki A, Leitner J, Abas L, Aerts M, Hoogewijs K, Baster P, De Groodt R, Lin Y-C, Storme V, Van de Peer Y, Beeckman T, Madder A, Devreese B, Luschnig C, Hilson P (2012) GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses. Dev Cell **22**: 678-685

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y (1999) *Oryza sativa PSK* gene encodes a precursor of phytosulfokine-α, a sulfated peptide growth factor found in plants. Proc Natl Acad Sci USA **96**: 13560-13565

Supplemental figures

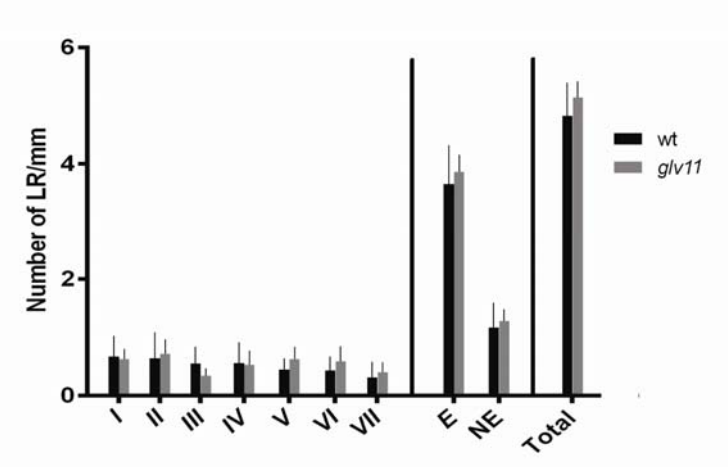


Figure S1: Distribution of LR developmental stages in control and *GLV11* loss-of-function lines.

Number of LRs per mm at different stages. For simplicity, results for one of two analyzed independent lines are presented (see Materials and Methods). Error bars show 95% confidence interval. Asterisks indicate significant differences compared to the wild type (P<0.001). I to VII, primordium stages; NE, nonemerged primordia; E, emerged LRs; total, total number of LRs.

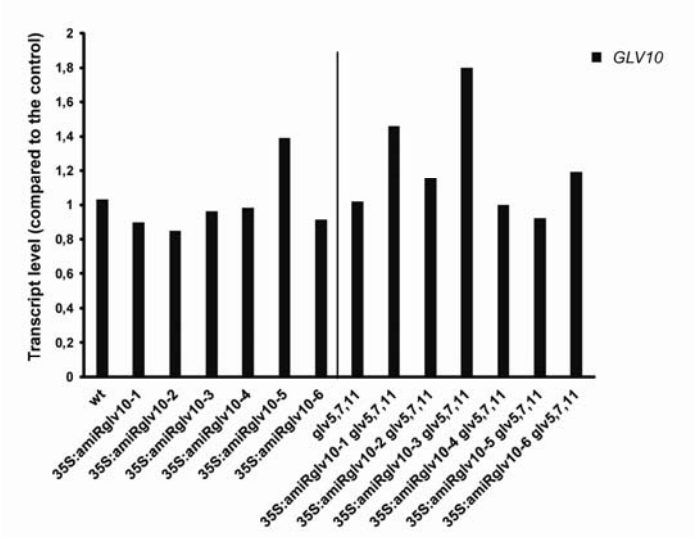


Figure S2: Relative transcript levels of the overexpression of *amiRglv10*.

Relative transcript levels of the overexpression of *amiRglv10* under the control of the strong *35S CaMV* promoter in Col-0 and *glv5,7,11* backgrounds compared to the wild type and *glv5,7,11*, respectively, measured by qRT-PCR analysis.

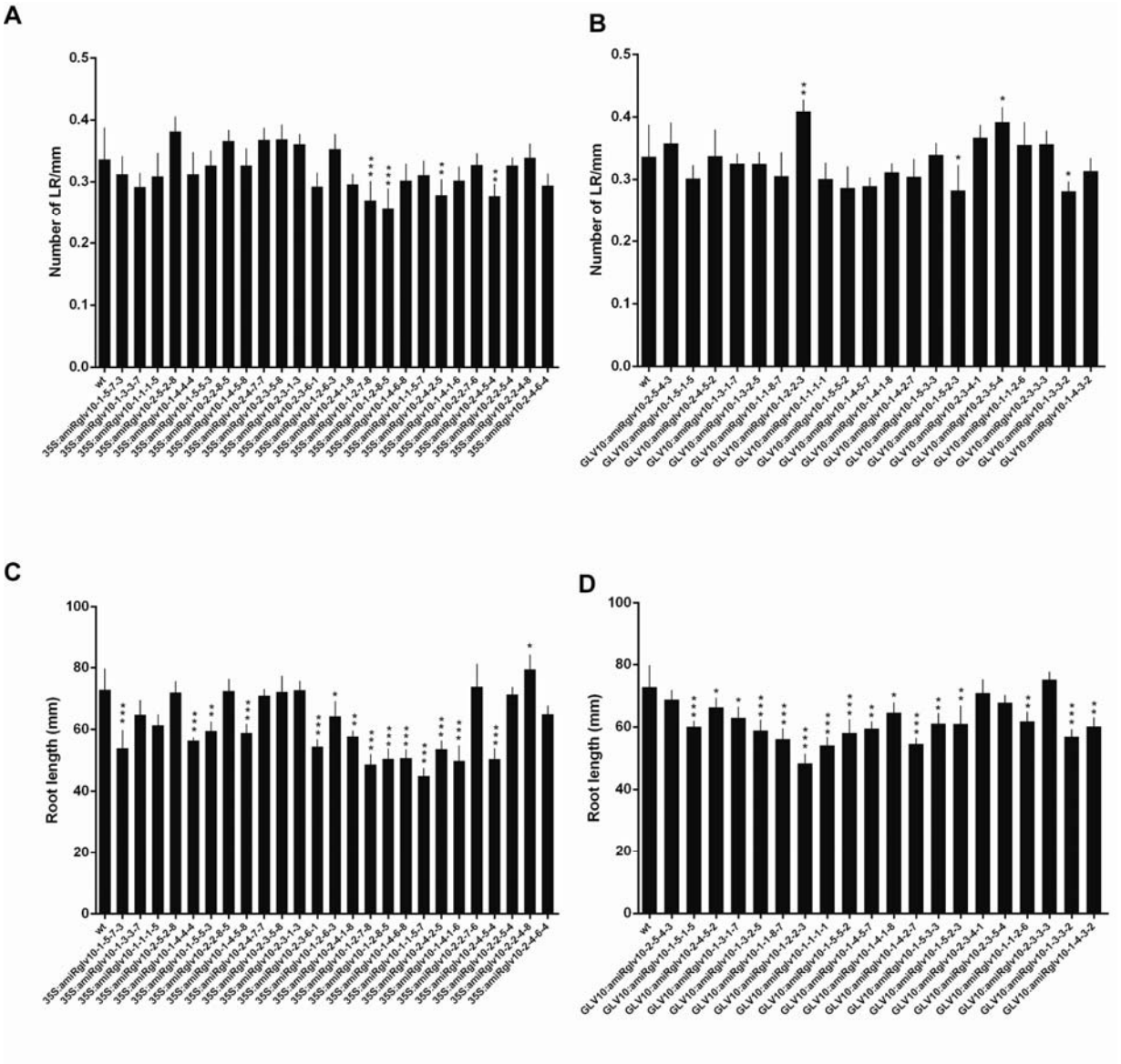


Figure S3-1: Root and LR development promoted by *GLV10* loss-of-function.

A and B, Emerged LR density in control and *GLV* loss-of-function lines (n=22-29). Results for one of three analyzed independent experiments are presented (see Materials and Methods). C and D, Quantification of root length (mm) in control and *GLV* loss-of-function lines (n=19-23). Root length was measured in 12-dag seedlings. For clarity, only one of three independent experiments is shown. Asterisks indicate significant differences compared to the wild type (P<0.001). *GLV10*-silenced independent mutant lines in **Col-0 background** driven by either the strong *35S CaMV* or *GLV10*-endogenous promoters.

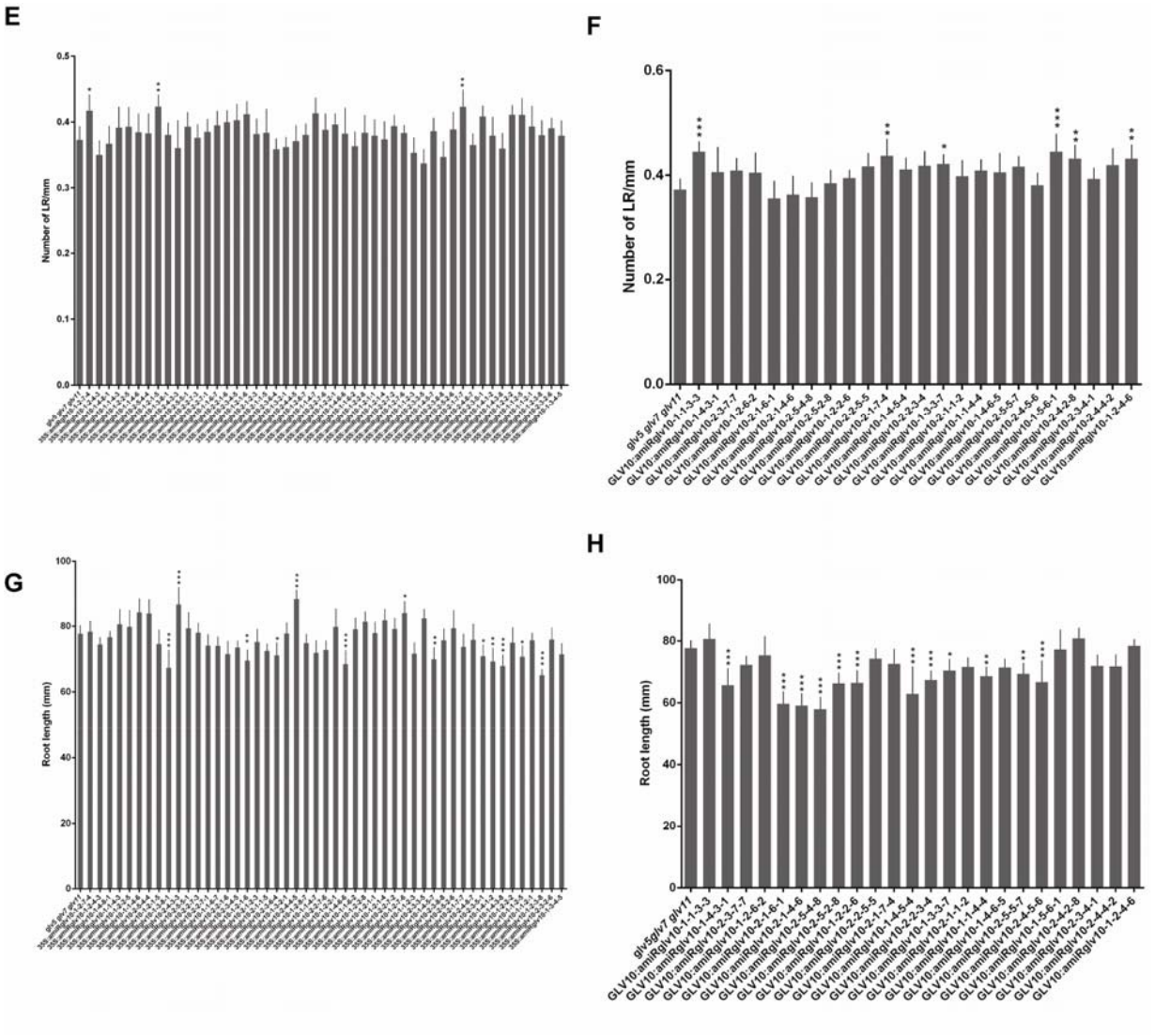


Figure S3-2: Root and LR development promoted by *GLV10* loss-of-function.

E and F, Emerged LR density in control and *GLV* loss-of-function lines (n=22-29). Results for one of three analyzed independent experiments are presented (see Materials and Methods). G and H, Quantification of root length (mm) in control and *GLV* loss-of-function lines (n=19-23). Root length was measured in 14-dag seedlings. For clarity, only one of three independent experiments is shown. Asterisks indicate significant differences compared to the wild type (P<0.001). *GLV10*-silenced independent mutant lines in the ***glv5glv7glv11*** **triple knockout** mutant background driven by either the strong *35S CaMV* or *GLV10*-endogenous promoters.

Table S1. List of a*miRglv10* genotypes

Original sequence (amiR target sequence)	Driven promoter	Background genotype	Genotype
Sequence 1: TTAATCTGCAAATAAGCGCGG	*GLV10*	Col-0	*GLV10:amiRglv10-1*
		glv5,7,11	*GLV10:amiRglv10-1 glv5,7,11*
	35S CaMV	Col-0	*35S:amiRglv10-1*
		glv5,7,11	*35S:amiRglv10-1 glv5,7,11*
Sequence 2: TCTAAGAGGATTATGAACCAT	*GLV10*	Col-0	*GLV10:amiRglv10-2*
		glv5,7,11	*GLV10:amiRglv10-2 glv5,7,11*
	35S CaMV	Col-0	*35S:amiRglv10-2*
		glv5,7,11	*35S:amiRglv10-2 glv5,7,11*

Table S2. List of the primers used to silence *GLV10* expression and qRT-PCR

Primer name	**Sequence 5′–3′**
I miR-s	gaTTAATCTGCAAATAAGCGCGGtctctcttttgtattcc
II miR-a	gaCCGCGCTTATTTGCAGATTAAtcaaagagaatcaatga
III miR*s	gaCCACGCTTATTTGGAGATTATtcacaggtcgtgatatg
IV miR*a	gaATAATCTCCAAATAAGCGTGGtctacatatatattcct
I miR-s	gaTCTAAGAGGATTATGAACCATtctctcttttgtattcc
II miR-a	gaATGGTTCATAATCCTCTTAGAtcaaagagaatcaatga
III miR*s	gaATAGTTCATAATCGTCTTAGTtcacaggtcgtgatatg
IV miR*a	gaACTAAGACGATTATGAACTATtctacatatatattcct
GLV10-Fw-qPCR	CACCATCGTCGTCCACCGCT
GLV10-Rev-qPCR	GGAGGCCGAGTGGACGGCTT

Chapter 5

A novel set of secretory peptides controlling root development

This chapter is in preparation for submission. SG, YCL, AF, TB, YVdP and PH designed the research. SG, YCL, BP and PH wrote the manuscript, with the help of AF, YVdP and TB. SG studied the transcription of small secretory peptides (SSP) genes in Arabidopsis tissues and characterized how selected SSPs alter root development (Figures 3 and 5). YCL created the algorithms enabling the search for genes coding for SSPs in plant genomes and the classification of related gene families, with the help of AF, YVdP and PH (Figures 1, 2 and 3). BP searched the pattern of expression of SSP genes in Arabidopsis transcriptome data. SG and YCL contributed equally to this work.

A novel set of secretory peptides controlling root development

Running head: Small peptides control root development

Sarieh Ghorbani[12]*, Yao-Cheng Lin[12]*, Boris Parizot[12], Ana Fernandez[12], Yves Van de Peer[12], Tom Beeckman[12], Pierre Hilson[1234]

[1]Department of Plant Systems Biology, VIB, B-9052 Ghent, Belgium

[2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium

[3]INRA, UMR1318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

[4]AgroParisTech, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

*These authors contributed equally to this work.

Corresponding author

Abstract

Plant genomes encode numerous small secretory peptides (SSPs) the functions of which remain yet to be explored. Based on structural features that characterize SSP families known to take part in postembryonic development, our comparative and inter-genomic analysis resulted in the identification of novel genes coding for oligopeptides potentially involved in cell-to-cell communication. Searching for putative regulators of root development, we selected, through meta-analysis, SSP genes that were expressed at specific stages and in specific cell types in the course of lateral root formation in Arabidopsis. As a first demonstration of their role in development, we showed that root growth and branching were altered by the application of synthetic peptides encoded by the respective SSP genes. Although further indepth research is required to unmask their precise role in root development, this two-step strategy combined of comparative genomics and rapid functional assays *in planta*, represents a simple approach to pinpoint factors potentially involved in non-cell-autonomous regulatory mechanisms. This approach can be implemented in different species and used for the study of a wide range of developmental programs.

Introduction

Plants are complex organisms that consist of distinct cell types organized in tissues. Separate plant organs as well as neighboring cells exchange a wide range of signals to coordinate development and to respond to environmental stimuli. The phytohormones that had initially been recognized to control plant growth and development are relatively simple chemicals. In recent years, peptides secreted into the apoplast by plant cells have also been identified as extracellular signals involved in various biological processes (Murphy et al., 2012). These bioactive molecules are referred, hereafter, as small secretory peptides (SSPs) (Matsubayashi, 2011b). Most SSPs are synthesized as preproproteins from which the signal sequence is cleaved upon targeting in the endoplasmic reticulum and further processed by successive proteolytic cleavages through the secretory pathway. Subclasses of cysteine-poor SSPs also undergo additional posttranslational modifications among which proline hydroxylation, hydroxyproline arabinosylation, and tyrosine sulfation have been documented (Matsubayashi, 2011a). All of them are involved in signaling pathways shown to control postembryonic growth and development (Murphy et al., 2012).

Because plants are sessile organisms, they have evolved a remarkable developmental plasticity in order to adapt to a wide range of ecological niches (Guyomarc'h, 2010). For example, a single embryonic root grows and branches to produce the entire root system through a finely coordinated developmental process that integrates endogenous and environmental cues. Multiple reports have shown that SSPs play an important role in meristem establishment, maintenance, cell division, lateral root (LR) initiation, development and emergence (recently reviewed in Somssich and Simon, 2012; Delay et al., 2013a).

In Arabidopsis, the formation of the LR primordium (LRP) has been described as eight successive stages, from the first asymmetric division of the pericycle founder cells to the LR emergence (Malamy and Benfey, 1997). The study of promoter-reporter constructs revealed that *GOLVEN* (*GLV*) genes are expressed differentially in specific cells from stage I until the establishment of the LR meristem (Fernandez et al., 2013a). Overproduction of GLV peptides resulted in a decreased number of LRs and perturbed the fine divisions in LRP (Meng et al., 2012; Fernandez et al., 2013a). Besides their known role in floral organ abscission, the INFLORESCENCE DEFICIENT IN ABCISSION (IDA) peptide, together with its receptors HAESA (HAE) and HAESA-Like 2 (HSL2), have recently been shown to be involved in LR

emergence (Kumpf et al., 2013). Moreover, a role in LR development has been proposed for the C-TERMINALLY ENCODED PEPTIDE 1 (CEP1) in Arabidopsis and *Medicago truncatula,* as demonstrated by the LR inhibition resulting from overexpression of *CEP1* or application of the peptide (Ohyama et al., 2008; Imin et al., 2013; Delay et al., 2013b). Finally, a regulatory module has been identified in which ETHYLENE-RESPONSIVE FACTOR 115 (ERF115), specifically expressed in the root quiescent center (QC), acts as a rate-limiting factor of cell division and is a direct activator of PSK5, a peptide of the phytosulfokine family, known to control cell division (Heyman et al., 2013).

Previous studies suggest that plant genomes contain more SSP genes than those identified until now and of which the function remains to be established (Lease and Walker, 2006; Silverstein et al., 2007; Lease and Walker, 2010). Indeed, the annotation of genes coding for SSPs is problematic because they harbor fewer characteristics of protein-coding sequences than larger genes and their similarity is restricted to domains coding for just a few amino acid residues. Therefore, bioinformatic pipelines relying on relatively simple sequence homology searches do not accurately predict SSP genes (Oelkers et al., 2008). Furthermore, hypothetical short open reading frames (ORFs) may arise by chance, albeit without function. Therefore, small ORFs are often under predicted or systematically removed in genome annotation projects, as was the case in early releases of the *Arabidopsis* genome. Alternatively, the detection of mature SSPs from crude plant tissue extracts is difficult because they are present at very low physiological concentrations (nanomolar range) and are generally masked by degradation products of larger and much more abundant proteins. Hence, it is likely that only a portion of the functional SSPs are known to date.

We propose an original method to identify unknown SSPs encoded in plant genomes without prior knowledge of their sequence by assuming that they share short conserved oligopeptide stretches. First, we detected patterns common to known plant SSPs with the aim to improve their recognition in genome sequences, including those of crop species. A blind sequence similarity search successfully identified well-characterized SSP families as well as unknown peptide-encoding genes, possibly involved in cell-to-cell communication. The consensus motifs resulting from this analysis corresponded to known SSP-derived bioactive peptides or defined potential unspecified families. We further investigated whether previously

uncharacterized SSPs might be involved in root development and showed that some of the corresponding genes were expressed in specific cell types and at particular stages of LR initiation. Finally, we demonstrated that the majority of the synthetic peptides matching conserved motifs in previously uncharacterized families strongly inhibit LR emergence.

Results

Identification of SSP genes in reference plant genomes

To validate our SSP identification algorithms, we chose as reference the preproprotein primary sequences of signaling peptides that share common structural features across plant gene families (identified first in *Arabidopsis thaliana* in most cases) involved in root development. They include: CEP, CLAVATA3 (CLV3/CLE), GOLVEN/ROOT GROWTH FACTOR/CLE-LIKE (GLV/RGF/CLEL), IDA, PSK, PLANT PEPTIDE CONTAINING SULFATED TYROSINE (PSY); and additional cysteine-rich peptides. In total, 216 Arabidopsis protein sequences were collected from these known secretory peptide families (Table 1 and Supplemental Table S1) as a benchmark set to test our *de novo* SSP detection strategy. Most of these short preproproteins contain an amino (N) -terminal signal peptide and a conserved carboxyl (C)-terminal end that is cleaved off to yield the mature signal. This latter sequence corresponds to the secreted bioactive portion of the peptide hormones that has been shown in multiple cases to act as a ligand of leucine-rich repeat-receptor-like kinase (LRR-RLK) membrane proteins (Butenko et al., 2009; Benkova and Hejatko, 2009; Murphy et al., 2012).

First, we built an exhaustive repertoire of C-terminally conserved domains across multiple plant species. Because the accuracy of gene models is crucial in this context, we only included species for which reliable genome annotations were available at the time this analysis was conducted: Arabidopsis, rice (*Oryza sativa*), poplar (*Populus trichocarpa*), grapevine (*Vitis vinifera*) and maize (*Zea mays*) (see Materials and Methods for details). The successive stages of our analytical pipeline aimed at identifying SSPs are explained here below and summarized in Figure 1.

Length - The average protein sequence length in the SSP benchmark set was 102 amino acids (Supplemental Table S1). The threshold of 200 amino acids was chosen as a conservative cutoff to exclude long protein sequences, resulting in 158,135 proteins selected from the predicted proteomes (including splice variants). About 24% of the predicted Arabidopsis proteins were shorter than 200 amino acids, yet the arbitrary protein sequence length cutoff removed only

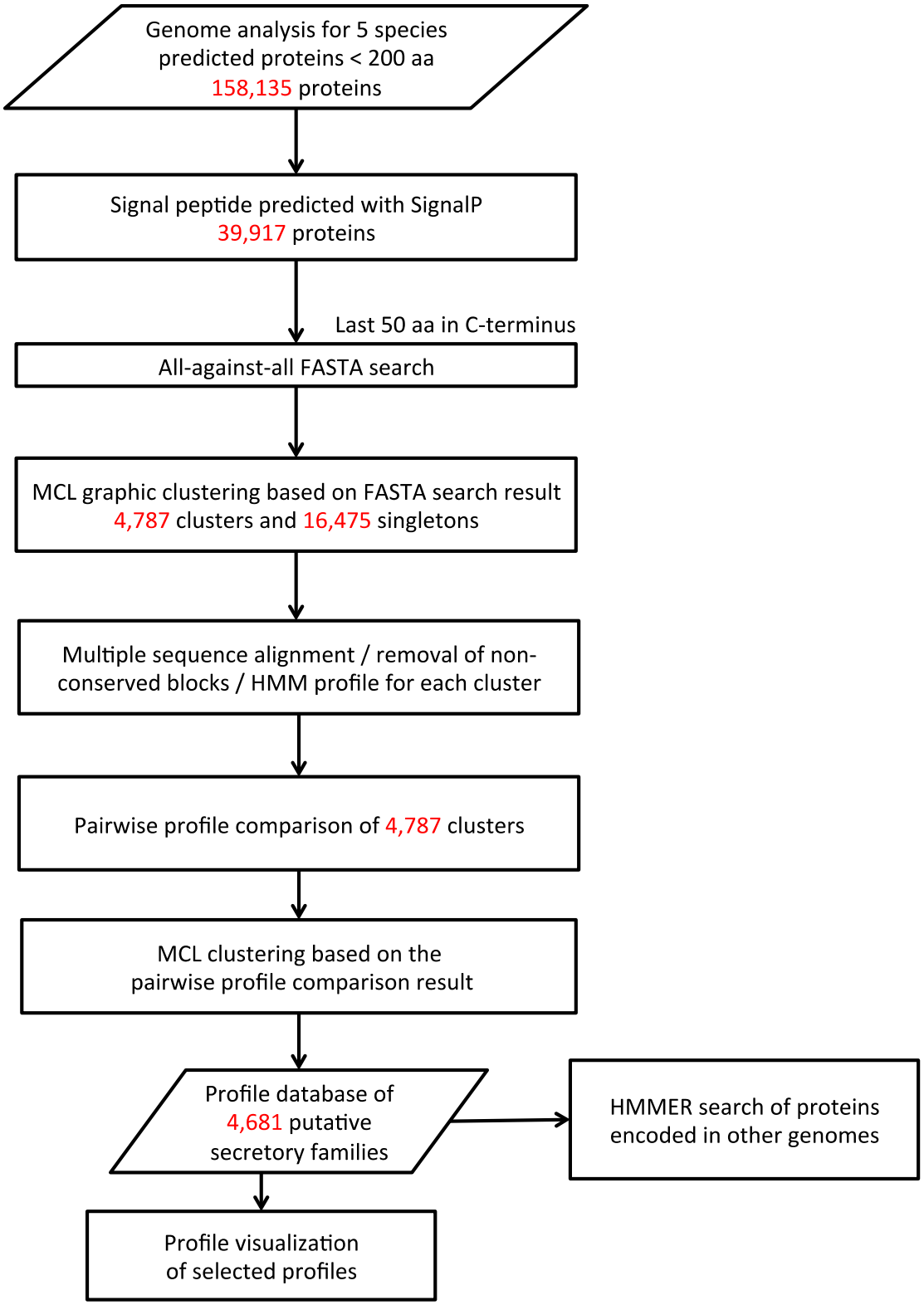


Figure 1. Flow chart of the pipeline for SSP family assembly. See Materials and Methods for details.

5 out of 216 secretory peptides (2.3%) from the benchmark dataset (CEP [At1G31670], At3G50610, gibberellic acid-stimulated in Arabidopsis [GASA; At5G14920], putative precursor for endogenous peptide elicitor [PROPEP; At1G17750] and At1G73080).

Secretion - 39,917 of these short proteins were predicted to contain an N-terminal hydrophobic region recognized as a cleavable signal sequence. However, not all characterized secretory signaling peptides carry such an identifiable sequence. Among the benchmark proteins, 40 (18.5%) did not contain a conventional signal peptide signal, which may also be partly explained by the arbitrary choice for the SignalP peptide identification parameters (Emanuelsson et al., 2007).

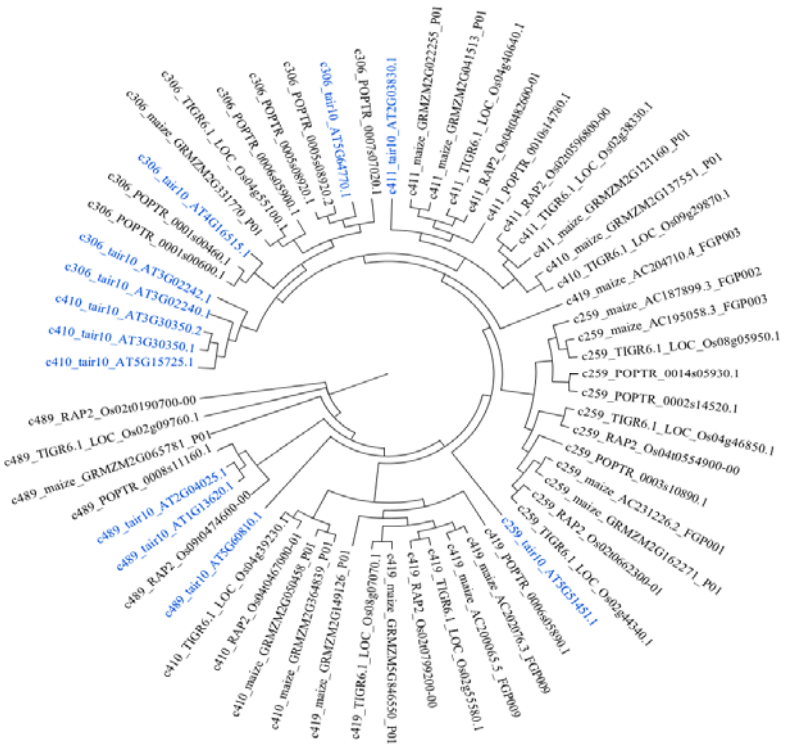
Conserved C-terminal motif - To reduce noise in sequence comparison, only the last 50 amino acids of the proteins were considered in the blind all-against-all FASTA sequence similarity search (Pearson, 2000). The first round of aggregation with the Markov Cluster Algorithm (MCL) grouped 23,442 proteins into 4,787 clusters and left out 16,475 proteins as singletons.

SSP family assembly

The candidate secretory peptides were further classified according to sequence homology by combining graphic clustering algorithms and pairwise profile comparisons (see Materials and Methods for details). To evaluate the performance of the clustering parameters, we examined the assembly of the known Arabidopsis CLV3/CLE and GLV/RGF/CLEL secretory signaling peptides. After the initial MCL clustering yielding 4,787 clusters, the 32 CLE Arabidopsis proteins were still scattered in seven clusters and the 11 Arabidopsis GLV proteins (including one splice variant) in five clusters.

The relationship between each cluster was then calculated via pairwise profile comparison and their higher-order relationship was determined with the MCL algorithm to link clusters into families. The MCL clustering based on the protein profiles markedly improved the resolution of secretory families. For example, the Arabidopsis CLE peptides clustered into two families corresponding to the subgroups involved in either root apical meristem (RAM) maintenance or vascular development (Kiyohara and Sawa, 2012), while the Arabidopsis GLV peptides were all grouped in a single family (Table 1) (Figure 2A).

A



B

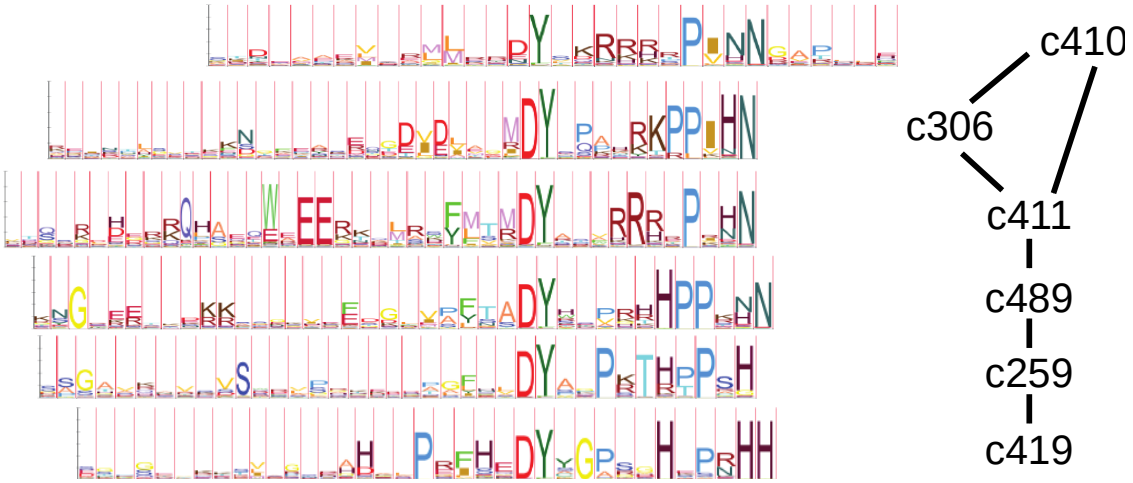


Figure 2. Reconstitution of the known GLV family after *de novo* global sequence comparison. A, Phylogenetic tree of the GLV family across six plants genome annotation. Cluster IDs from the first MCL clustering are indicated in the first prefix of each protein sequence and the species ID (data source) corresponds to the second prefix. Known Arabidopsis GLV peptides are highlighted in blue. TAIR10: *A. thaliana* TAIR10; RAP2: *O. sativa* RAP-DB, IRGSPbuild5; TIGR6.1: *O. sativa* MSU 6.1; PORTR: *P. trichocarpa* JGI v156; vitis: *V. vinifera*, Genoscope v1 and maize: *Z. mays* ZmB73_5a. B, GLV/RGF/CLEL cluster relationships. Black lines indicate the connectivity of each GLV cluster based on the pairwise profile comparison.

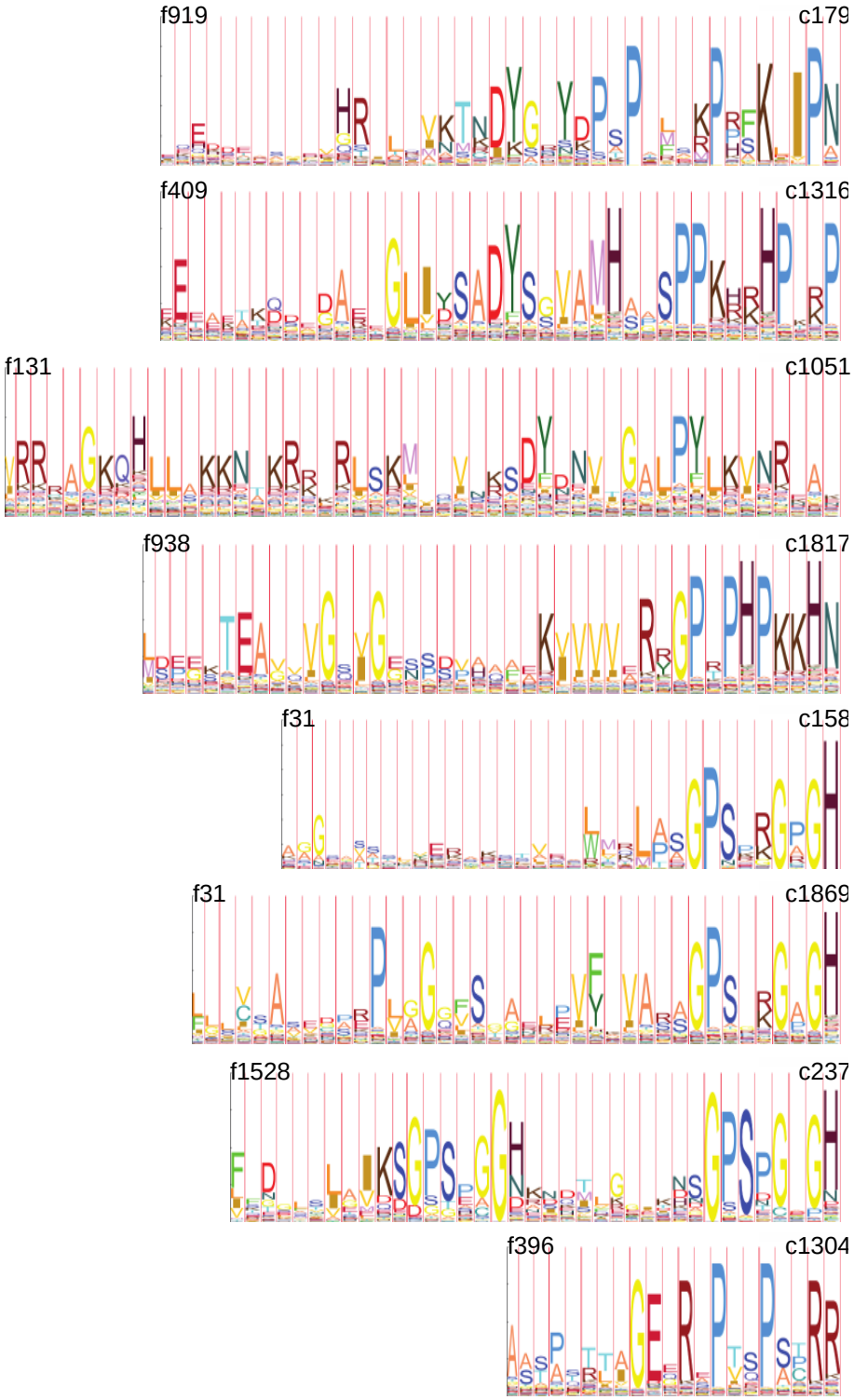


Figure 3. Conserved SSP C-terminal sequences. Consensus sequence representation for selected previously uncharacterized families. Conserved protein residues are higher in the HMM profile.

As expected, the topology of the cluster connectivity network built with the predicted proteins selected from the five species resembles the phylogenetic relationships between peptides in a given family: close sequences according to the phylogenetic tree tend to group together in the same cluster or in neighboring clusters (Figure 2B). The consistency between the two independent sequence clustering approaches indicates that our analytical pipeline can reconstruct the secretory peptide families without prior knowledge of the sequence information.

Furthermore, the manual curation of previously unreported consensus sequences revealed conspicuous patterns commonly observed in known signaling peptide families. For example, a tyrosine residue was found at the N-side of the conserved motif in multiple groups (for example f131, f409, f919; Figure 3). This tyrosine is known to be sulfated in the GLV, PSK and PSY mature signaling peptides, where it is also preceded by an aspartic acid residue. Its presence and its posttranslational modification are crucial for bioactivity (Komori et al., 2009; Matsuzaki et al., 2010; Whitford et al., 2012). The conserved motifs often end at or very near the last C-terminal residue of the precursor protein and contain one or several proline residues that might act as hinges when the peptide ligand binds to its receptor (Figure 3). Together, these observations indicate that our global *de novo* sequence search method provides valuable hints about potential unrecognized *bona fide* SSPs.

SSP gene regulation in the course of root development

Considering the established role of several secretory peptides in root development, we examined how SSP genes were expressed during LR formation in Arabidopsis. Our aim was to test whether the spatiotemporal specificity of their transcription pattern could be a valuable predictor for their possible involvement in root development. To this end, we analyzed the SSP transcript levels in transcriptome experiments addressing early aspects of the LR initiation, taking place in the pericycle associated with the xylem poles and depending on SOLITARY ROOT/INDOLE-3-ACETIC ACID14 (SLR/IAA14) mediated auxin signaling cascade. Three datasets follow the transcriptional regulations occurring during the induction of the LR initiation upon treatment (1) with auxin and depending on SLR/IAA14 (Vanneste et al., 2005); (2) with auxin and naxillin, a non-auxin-like lateral root inducing molecule (De Rybel et al., 2012); (3) with auxin, tracking specifically changes in the pericycle cells at the xylem pole (De Smet et al., 2008). Two other datasets address the gene spatial expression pattern: (4) differential between the

pericycle cells at the xylem or phloem pole (Parizot et al., 2012); (5) specificity in the LRP, either in the entire pericycle, or in one of its subpopulations (xylem or phloem pole) (Brady et al., 2007). The last dataset (6) focuses on the temporal expression pattern in phase or anti-phase with the auxin transcriptional response marker DR5, in the basal meristem (Moreno-Risueno et al., 2010).

First, we searched the transcriptomics data for patterns associated with known SSP gene families (Table 1). A portion of the SSP sequences are not represented on the Affymetrix ATH1 microarray (65 out of 148; 44%). Out of the 83 known SSP genes with a corresponding probeset, 50% had a specific spatiotemporal expression pattern in a least one of the analyzed experiments (fold change [FC] $\geq$ 1.5, P $\leq$ 0.01; for additional information, see Materials and Methods) (Supplemental Table S2). This observation sugggests that many more secretory peptides might be involved in apoplastic signaling during LR initiation than previously recognized.

We extended our analysis to genes belonging to uncharacterized SSP families and represented on ATH1 microarray. In four out of seven cases analyzed, significant changes in expression were detected (Table 2). At4G37295, At4G34600, and At4G37290 are induced in the xylem pole pericycle upon auxin treatment and depend on the IAA14/SLR pathway. At4G37295 and At4G37290 are also induced upon naxillin treatment. At4G37295 is specifically expressed in the LRP. At4G28460 and At1G49800 are in phase with the oscillating auxin response observed in the basal meristem with the DR5 marker, and the expression of At4G28460 is also higher in the phloem pole pericycle than that in the xylem pole pericycle. In conclusion, the expression of a large fraction of SSP-encoding genes is regulated during LR initiation, whether they have been recognized previously as involved in development or not.

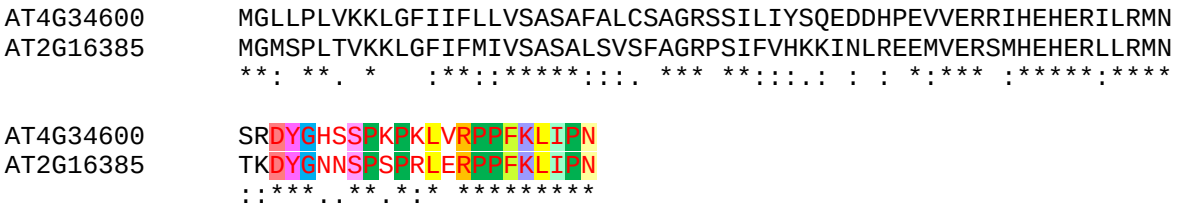
SSP functional analysis

The activity of SSPs can be tested by the application of chemically synthetized peptides on plant tissues because the response they induce often copies the cognate genetic gain-of-function phenotypes, as shown in Arabidopsis roots (Fiers et al., 2006; Matsuzaki et al., 2010; Whitford et al., 2012; Fernandez et al., 2013a). Such experiments demonstrated that the bioactive portion of the SSP preproproteins is encoded in their C-terminal conserved sequences.

To investigate the potential role of uncharacterized SSPs, we grew seedlings on agar medium supplemented with synthetic peptides corresponding to conserved C-terminal stretches (Figure 4; Supplemental Table S3). Whereas synthetic SSPs, including members of the CLV3/CLE and GLV/RGF/CLEL families, are active at nanomolar concentrations (Murphy et al., 2012), the absence of certain posttranslational modifications in synthetic copies compared to native peptides has been shown to reduce bioactivity (Seitz, 2000; Matsubayashi, 2011a; Shinohara and Matsubayashi, 2013). To avoid false negative results, we applied micromolar concentrations of synthetic peptides, as commonly reported in such experiments.

New family 1 (f31)

New family 2 (f919)



New family 3 (f1528)

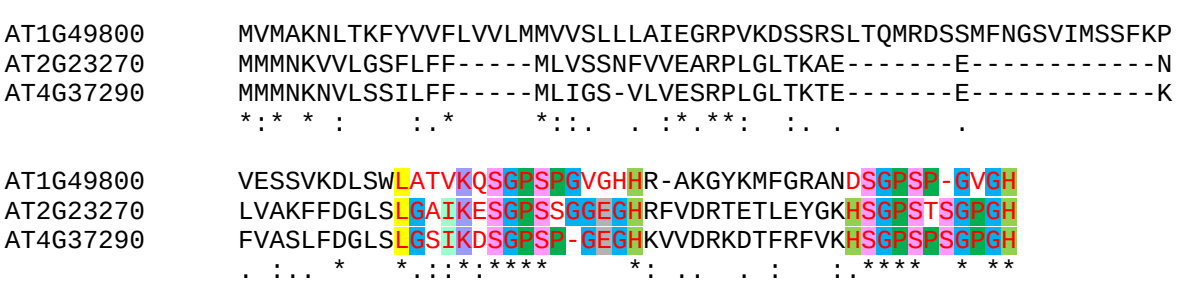


Figure 4. Primary sequence alignment of Arabidopsis SSPs tested in root development assays. The multiple sequence alignment generated with ClustalW2.

We compared the number of LRs and the primary root length between control seedlings and seedlings treated with 1 or 10 µM of peptides. Peptides (Pep) from the new families 1 and 2 decreased the number of emerged LRs, with a particular mention for Pep2-2 (At4G34600) resulting in a 70% decrease, compared to control untreated seedlings (Figure 5A). In all cases, the effect was stronger or only detectable at 10 µM. Furthermore, plantlets treated with 10 µM of Pep1-2 (At4G37295) were pale and arrested in growth. From the new family 3, only Pep3-2-2 (At2G23270) and Pep3-3-2 (At4G37290) induced significant differences compared to control untreated plants (Figure 5A). Comparison of the primary root length of seedlings grown in the same conditions revealed no or only slight differences with a reduction after treatment with Pep1-1 and Pep2-2 at 1 µM and Pep2-1 and Pep3-2-1 at 10 µM (Figure 5B).

To confirm the role of the corresponding uncharacterized SPP genes in LR development, we quantified their transcriptional changes in the LR-inducible system. In this experimental set up, the first formative divisions are prevented by the auxin transport inhibitor *N*-1-naphthylphthalamic acid (NPA). Later, upon auxin (1-naphthaleneacetic acid, NAA) treatment, cells in the pericycle layer engage actively and synchronously in division (Himanen et al., 2002). Reverse-transcription polymerase chain reaction (qRT-PCR) analysis showed very specific transcription patterns for some candidates (Figure 5C).

Expression of the genes analyzed increased after both 2 and 6 h for AT4G37295, AT5G43066, AT4G37290, and AT2G16385, but continuously decreased for AT4G28460 and AT4G34600. Expression level of AT3G06090 and AT2G23270 decreased after 2 h and increased after 6 h, while AT1G49800 had the opposite pattern. These changes are in accordance with the transcriptome data and strongly indicate that the tested genes act at specific stages and in specific tissues during root development, including LR initiation (Ohyama et al., 2008; Fernandez et al., 2013a).

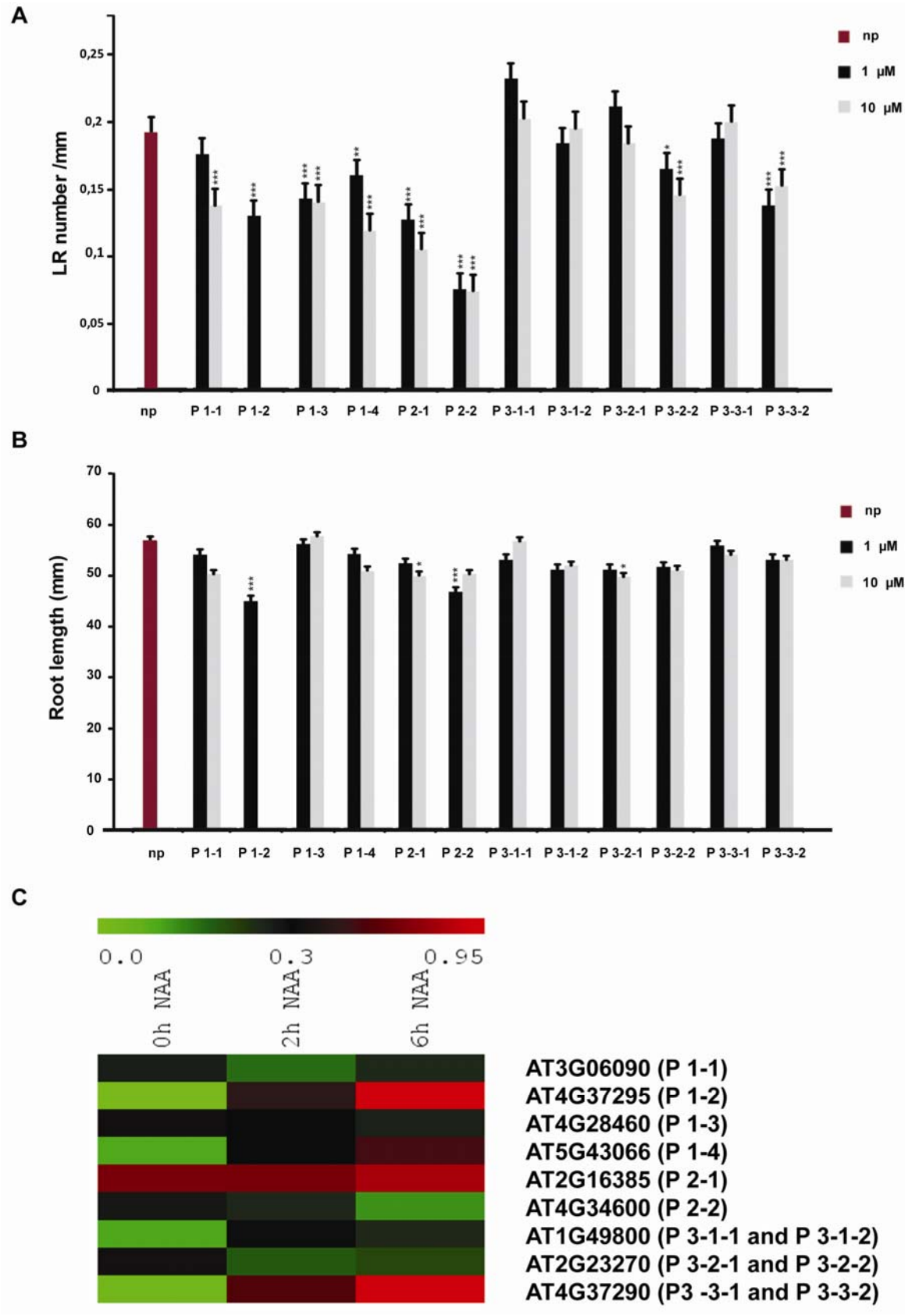

A
LR number /mm
0,25
0,2
0,15
0,1
0,05
0
np
1 µM
10 µM
np
P 1-1
P 1-2
P 1-3
P 1-4
P 2-1
P 2-2
P 3-1-1
P 3-1-2
P 3-2-1
P 3-2-2
P 3-3-1
P 3-3-2
B
Root lemgth (mm)
70
60
50
40
30
20
10
0
np
1 µM
10 µM
C
0.0
0.3
0.95
0h NAA
2h NAA
6h NAA
AT3G06090 (P 1-1)
AT4G37295 (P 1-2)
AT4G28460 (P 1-3)
AT5G43066 (P 1-4)
AT2G16385 (P 2-1)
AT4G34600 (P 2-2)
AT1G49800 (P 3-1-1 and P 3-1-2)
AT2G23270 (P 3-2-1 and P 3-2-2)
AT4G37290 (P3 -3-1 and P 3-3-2)

Figure 5. Root-related phenotypes induced by extra identified SSPs. (A) Number of emerged LRs per unit length (mm) (n=20-37). (B) Primary root length (n=19-44). Ten-dag seedlings were compared with controls after treatment with the indicated peptides. Error bars represent the 95% confidence interval. Asterisks mark significant differences: * $P < 0.05$; ** $P < 0.005$, *** $P < 0.001$. Data were pooled from two independent replicates. (C) Induction of SSP gene transcription by auxin. Seedlings were treated with 1 µM NAA for the indicated time points. Fold-changes were measured after qRT-PCR analysis of root tissues. Data were pooled from two independent replicates.

Discussion

A bottleneck in the study of the function and relevance of signaling peptides for plant growth and development has been the identification of the encoding genes. Whereas the sequencing of different plant genomes has led to the prediction of numerous small genes, some of which potentially encoding signaling peptides, the identification of conserved families via comparative genomics is difficult, because their bioactive domains are restricted to just a few amino acids.

Unlike previous studies solely relying on the SSP information embedded in the Arabidopsis genome annotation (Lease and Walker, 2006; Silverstein et al., 2007), our *de novo* comparative genomics approach takes advantage of genome annotation in other species without a prior knowledge of the SSP sequence information and classifies SSP families with high resolution. The presence of multiple plant species in the pipeline increases the sensitivity to separate large SSP families into multiple smaller groups. The subsequent profile comparison improves the clustering specificity. Previously separated small groups may be linked into large families based on the shared sequence profiles. This bioinformatic approach leads to a classification that can be updated rapidly and regularly as information accrues. New found consensus motifs can serve as functional domain hallmarks to search for small missed genes, either in assembled genome sequences or in shorter RNA-sequence reads.

We were able to point out the potential involvement of the SPPs in the process of LR development through the analysis of transcriptome data (Parizot et al., 2010). This method has already led to the discovery of several genes proven to be involved in LR development in follow up genetic studies (GATA23, De Rybel et al., 2010; E2Fa, Berckmans et al., 2011; PdBG1, Benitez-Alfonso et al., 2013; totipotency genes, Chupeau et al., 2013; PLT3, Zhang et al., 2013; PDCB1, Maule et al., 2013).

After identifying new candidate SSP families and showing that they include genes with specific expression patterns during LR initiation, we showed that the majority of conserved peptides tested altered the growth of Arabidopsis roots when applied exogenously. This type of assays is a cheap, easy and rapid first step towards the classification of non cell-autonomous factors involved in development and can be adapted to a wide range of processes.

However, the refined understanding of the SSP function requires additional studies to avoid the pitfalls of gain-of-function phenotypes: nonphysiological concentrations of signal molecules may create artifacts, for example by hijacking downstream pathways of related, but distinct, peptide signal(s); exogenous applications are nondirectional, whereas SSP genes are often expressed in very specific cell types, as again demonstrated here. Nevertheless, our results indicate that the successive combination of SSP gene annotation, expression studies and *in vivo* peptide assays is a useful approach to start speedily probing the complexity of the extracellular signaling networks that drive plant tissue growth and development.

MATERIALS AND METHODS

Selection of short proteins with signal peptide

As the *de novo* detection of secretory peptides is sensitive to the quality of the gene models, we selected six sequenced plant genomes with consistently improved annotations: *Arabidopsis thaliana* (TAIR10), rice (*Oryza sativa*; IRGSPbuild5 and MSU6.1) (Ouyang et al., 2007; Rice Annotation Project, 2008); poplar (*Populus trichocarpa*; JGI v156) (Tuskan et al., 2006); grapevine (*Vitis vinifera*; Genoscope v1) (Jaillon et al., 2007) and maize (*Zea mays*; ZmB73_5a) (Schnable et al., 2009). For all five species, genome annotations had been updated at least once after their initial release at the time this analysis was conducted, thus providing quality curated data. Two rice genome annotations were processed because their annotation of small predicted proteins was complementary. Only protein sequences of less than 200 amino acids in length were kept for further analysis. We searched for the presence of the signal peptide in the amino terminal domain with the SignalP v3.0 software (Bendtsen et al., 2004). The signal peptide was predicted with the neural network or HMM profile.

De novo conserved secretory motif detection

The last 50 amino acids from the candidate secretory peptides were searched against each other by the FASTA program (Pearson, 2000) with the BLOSUM50 scoring matrix to detect mildly related sequences. Second, the all-against-all FASTA search results were subjected to the Markov Clustering algorithm (MCL version 09-308, inflation value 1.5) (Enright et al., 2002) to identify the sequences into clusters based on the e-value. We paid special attention to the inflation point in the MCL algorithm because it controls the connectivity between related protein subgroups and the main challenge in the delineation of secretory peptide families is the weak sequence similarity between members. Third, sequences in each cluster were aligned by the multiple alignment program MUSCLE (Edgar, 2004); non-aligned gaps and non-conserved positions in the multiple alignment were removed based on the BLOSUM62 scoring matrix. Fourth, based on the remaining conserved region, each cluster was represented by a hidden Markov model (HMM) profile with hmmbuild and hmmcalibrate from the HMMER (v2.3.2) package (http://hmmer.wustl.edu/). Fifth, singleton sequences that did not cluster in the previous MCL clustering were searched (hmmersearch) against the HMM profiles to identify the most closely related clusters. When an additional sequence was identified in a cluster, this new

sequence was combined with the preexisting ones in that cluster, and the procedure was reinitiated from step three. We considered the search for a cluster completed once no sequence could be added to it.

The HMM profile of each cluster was compared against all HMM profiles by the Profile Comparer (PRC) (Madera, 2008). Then, the higher-order relationship of the clusters was determined with the MCL algorithm based on the e-values calculated with PRC. To inspect the shared conserved motif of candidate secretory cluster pairs, 'LogoMat-P' (Schuster-Bockler and Bateman, 2005) was applied to generate the pairwise HMM logos. A group of clusters linked by the PRC program was considered as one putative secretory family.

Microarray data normalization and compendium analysis

Transcriptome datasets were retrieved as Gene Expression Omnibus (GEO) accessions: GDS1515 (Vanneste et al., 2005), GSE42896 (De Rybel et al., 2012), GSE6349 (De Smet et al., 2008), and GSE8934 (Parizot et al., 2012) for the phloem and the xylem pole pericycle expression files. The full pericycle expression data, based on the J2661 Arabidopsis marker line, were a kind gift (J2661, Levesque et al., 2006). Array data were normalized with the robust multiarray average algorithm (Irizarry et al., 2003) and the absolute values, FC and pairwise P-values were determined with the affylmGUI R package (Smyth, 2004) without adjustment. Two-factor ANOVA P-values were computed with the MultiExperiment Viewer (http://www.tm4.org/mev/). Affymetrix probe sets were assigned to AGI gene ID according to the "affy_ATH1_array_elements-2010-12-20.txt" file from TAIR (www.arabidopsis.org). Ambiguously assigned genes (multiple gene identifiers for one probe set) and microarray controls were discarded. Genes were considered significantly regulated in specific experiments when the following criteria were fulfilled: absolute FC $\geq$ 1.5, P $\leq$ 0.01 for at least one of the pairwise comparisons (0-2, 2-6, 0-6 h) upon LR induction in the control plants, and a two-factor ANOVA P$\leq$ 0.01 for the interaction between treatment and genotype (Vanneste et al., 2005); absolute FC $\geq$ 1.5, P $\leq$ 0.01 for at least one of the pairwise comparison (0-2, 2-6, 0-6 h) for both compounds (NAA and naxillin) during the time course upon the LR induction system (De Rybel et al., 2012); absolute FC $\geq$ 1.5, P $\leq$ 0.01 for at least one of the pairwise comparison (0-2, 2-6, 0-6 h) during the time course upon LR initiation in the sorted pericycle cells (De Smet et al., 2008); absolute

FC ≥ 1.5, P ≤ 0.01 for at least one of the pairwise comparison (xylem pole pericycle vs. phloem pole pericycle, xylem pole pericycle vs. full pericycle, full pericycle vs. phloem pole pericycle) and similar positive or negative sign for all the pairwise comparisons (Parizot et al., 2012). Additionally, a radial layer specificity was determined as described in the respective publication of Brady et al. (2007) and a gene was tagged when specifically expressed in the xylem or phloem pericycle pole, or in the primordium. A radial layer specificity was determined as described (Brady et al., 2007) (Supplemental Table 2) and a gene was tagged when specifically expressed in the xylem or phloem pericycle pole, or in the primordium. Furthermore, an oscillation cluster association was determined as described (Moreno-Risueno et al., 2010) (Supplemental Table 1) and a gene was tagged when expressed in phase or antiphase with DR5 oscillation (Parizot et al., 2010).

Plant material and growth conditions

All experiments were conducted with wild-type *Arabidopsis thaliana* (L.) Heyhn., accession Columbia-0 (Col-0). Seeds were surface-sterilized and sown on half strength Murashige and Skoog (MS) medium (Duchefa Biochemie B.V.) complemented with 1% (w/v) agarose and 1.5% (w/v) sucrose at pH 5.8. Seeds were stratified for at least 2 days at 4°C. Seedlings were germinated in illuminated growth chambers under a 16-h light/8-h dark cycle (100 µmol $m^{-2}s^{-1}$) at 21°C. NPA and NAA treatments and transcript level assays were as described (Himanen et al. 2002).

Gene expression analysis

Total RNA from roots 5 days after germination (dag) was isolated with TRIzol reagent (Invitrogen), followed by treatment with RNase-free DNase I (Qiagen) according to the manufacturer's instructions. The cDNA was prepared with the iScript™cDNA Synthesis Kit (Bio-Rad) from 1 µg of total RNA and 1:10 dilutions of total cDNA were used as template for quantitative RT-PCR. Means of samples were compared with two-way analysis of variance (ANOVA) (GraphPad Prism; V6.00, GraphPad Software).

Acknowledgement

This work was supported by the Integrated Project AGRON-OMICS, in the Sixth Framework Programme of the European Commission (LSHG-CT-2006-037704) and the Interuniversity Attraction Poles Programme (IAP P7/29), initiated by the Belgian Science Policy Office. We thank Martine De Cook for help in the preparation of the manuscript.

Supplementary Information

Supplemental Information contains one figure and five tables and can be found online at

References

Amano Y, Tsubouchi H, Shinohara H, Ogawa M, Matsubayashi Y (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. Proc Natl Acad Sci USA **104**, 18333-18338.

Bendtsen JD, Nielsen H, von Heijne G, and Brunak S (2004). Improved prediction of signal peptides: SignalP 3.0. J Mol Biol **340,** 783-795.

Benitez-Alfonso Y, Faulkner C, Pendle A, Miyashima S, Helariutta Y, and Maule A (2013). Symplastic intercellular connectivity regulates lateral root patterning. Dev Cell **26,** 136-147.

Benkova E, and Hejatko J (2009). Hormone interactions at the root apical meristem. Plant Mol Biol **69,** 383-396.

Berckmans B, Vassileva V, Schmid SPC, Maes S, Parizot B, Naramoto S, Magyar Z, Kamei CLA, Koncz C, Bogre L, Persiau G, De Jaeger G, Friml J, Simon R, Beeckman T, and De Veylder L (2011). Auxin-Dependent Cell Cycle Reactivation through Transcriptional Regulation of Arabidopsis E2Fa by Lateral Organ Boundary Proteins. Plant Cell **23,** 3671-3683.

Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. Science **318**: 801-806

Chupeau MC, Granier F, Pichon O, Renou JP, Gaudin V, and Chupeau Y (2013). Characterization of the Early Events Leading to Totipotency in an Arabidopsis Protoplast Liquid Culture by Temporal Transcript Profiling. Plant Cell **25,** 2444-2463.

De Rybel B, Audenaert D, Xuan W, Overvoorde P, Strader LC, Kepinski S, Hoye R, Brisbois R, Parizot B, Vanneste S, Liu X, Gilday A, Graham IA, Nguyen L, Jansen L, Njo MF, Inze D, Bartel B, and Beeckman T (2012). A role for the root cap in root branching revealed by the non-auxin probe naxillin. Nature chemical biology **8,** 798-805.

De Rybel B, Vassileva V, Parizot B, Demeulenaere M, Grunewald W, Audenaert D, Van Campenhout J, Overvoorde P, Jansen L, Vanneste S, Moller B, Wilson M, Holman T, Van Isterdael G, Brunoud G, Vuylsteke M, Vernoux T, De Veylder L, Inze D, Weijers D, Bennett MJ, and Beeckman T (2010). A Novel Aux/IAA28 Signaling Cascade Activates GATA23-Dependent Specification of Lateral Root Founder Cell Identity. Current Biology **20,** 1697-1706.

De Smet I, Vassileva V, De Rybel B, Levesque MP, Grunewald W, Van Damme D, Van Noorden G, Naudts M, Van Isterdael G, De Clercq R, Wang JY, Meuli N, Vanneste S, Friml J, Hilson P, Jurgens G, Ingram GC, Inze D, Benfey PN, and Beeckman T (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the Arabidopsis root. Science **322,** 594-597.

Delay C, Imin N, and Djordjevic MA (2013a). Regulation of root development by small signaling peptides. Frontiers in plant science **4,** 352.

Delay C, Imin N, Djordjevic MA (2013b) *CEP* genes regulate root and shoot development in response to environmental cues and are specific to seed plants. J Exp Bot **64**, 5383-94.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32,** 1792-1797.

Emanuelsson O, Brunak S, von Heijne G, and Nielsen H (2007). Locating proteins in the cell using TargetP, SignalP and related tools. Nature protocols **2,** 953-971.

Enright AJ, Van Dongen S, and Ouzounis CA (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res **30,** 1575-1584.

Fernandez A, Drozdzecki A, Hoogewijs K, Nguyen A, Beeckman T, Madder A, and Hilson P (2013a). Transcriptional and Functional Classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-Like Signaling Peptides Reveals Their Role in Lateral Root and Hair Formation. Plant Physiol **161,** 954-970.

Fiers M, Golemiec E, van der Schors R, van der Geest L, Li KW, Stiekema WJ, and Liu CM (2006). The CLAVATA3/ESR motif of CLAVATA3 is functionally independent from the nonconserved flanking sequences. Plant Physiol **141,** 1284-1292.

Guyomarc'h SL, Mikaël l; and Laplaze L (2010). Lateral/Secondary Roots. in: Encyclopedia of Life Sciences (ELS). John Wiely & Sons, Ltd: Chichester.

Himanen K, Boucheron E, Vanneste S, de Almeida Engler J, Inze D, and Beeckman T (2002). Auxin-mediated cell cycle activation during early lateral root initiation. Plant Cell **14,** 2339-2351.

Imin N, Mohd-Radzman NA, Ogilvie HA, and Djordjevic MA (2013). The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula. J Exp Bot **64,** 5395-5409.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, and Speed TP (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4,** 249-264.

J Maule, A, Gaudioso-Pedraza R, and Benitez-Alfonso Y (2013). Callose deposition and symplastic connectivity are regulated prior to lateral root emergence. Communicative & Integrative Biology **6,** e26531.

Jaillon O, et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature **449,** 463-467.

Kumpf RP, Shi CL, Larrieu A, Sto IM, Butenko MA, Peret B, Riiser, ES, Bennett MJ, and Aalen RB (2013). Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence. P Natl Acad Sci USA **110,** 5235-5240.

Lease KA, and Walker JC (2006). The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol **142,** 831-838.

Lease KA, and Walker JC (2010). Bioinformatic identification of plant peptides. Methods in molecular biology **615,** 375-383.

Levesque MP, Vernoux T, Busch W, Cui H, Wang JY, Blilou I, Hassan H, Nakajima K, Matsumoto N, Lohmann JU, Scheres B, and Benfey PN (2006). Whole-Genome Analysis of the SHORT-ROOT Developmental Pathway in Arabidopsis. PLoS biology **4,** e143.

Madera M (2008). Profile Comparer: a program for scoring and aligning profile hidden Markov models. Bioinformatics **24,** 2630-2631.

Matsubayashi Y (2011a). Post-translational modifications in secreted peptide hormones in plants. Plant & cell physiology **52,** 5-13.

Matsubayashi Y (2011b). Small post-translationally modified Peptide signals in Arabidopsis. Arabidopsis Book 9, e0150.

Matsuzaki Y, Ogawa-Ohnishi M, Mori A, and Matsubayashi Y (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. Science **329,** 1065-1067.

Meng L, Buchanan BB, Feldman LJ, and Luan S (2012). CLE-like (CLEL) peptides control the pattern of root growth and lateral root development in Arabidopsis. Proc Natl Acad Sci USA **109,** 1760-1765.

Moreno-Risueno MA, Van Norman JM, Moreno A, Zhang JY, Ahnert SE, and Benfey PN (2010). Oscillating Gene Expression Determines Competence for Periodic Arabidopsis Root Branching. Science **329,** 1306-1311.

Murphy E, Smith S, and De Smet I (2012). Small Signaling Peptides in Arabidopsis Development: How Cells Communicate Over a Short Distance. Plant Cell **24,** 3198-3217.

Oelkers K, Goffard N, Weiller GF, Gresshoff PM, Mathesius U, and Frickey T (2008). Bioinformatic analysis of the CLE signaling peptide family. Bmc Plant Biol **8,** 1.

Ohyama K, Ogawa M, and Matsubayashi Y (2008). Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. The Plant Journal **55,** 152-160.

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, Orvis J, Haas B, Wortman J, and Buell CR (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res **35,** D883-887.

Parizot B, De Rybel B, and Beeckman T (2010). VisuaLRTC: a new view on lateral root initiation by combining specific transcriptome data sets. Plant Physiol **153,** 34-40.

Parizot B, Roberts I, Raes J, Beeckman T, and De Smet I (2012). In silico analyses of pericycle cell populations reinforce their relation with associated vasculature in Arabidopsis. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **367,** 1479-1488.

Pearson WR (2000). Flexible sequence similarity searching with the FASTA3 program package. Methods in molecular biology **132,** 185-219.

Rice Annotation Project, Tanaka T, et al. (2008). The Rice Annotation Project Database (RAP-DB): 2008 update. Nucleic Acids Res **36,** D1028-1033.

Roxrud I, Lid SE, Fletcher JC, Schmidt ED, Opsahl-Sorteberg HG (2007). GASA4, one of the 14-member Arabidopsis GASA family of small polypeptides, regulates flowering and seed development. Plant Cell Physiol. **48,**471-83.

Schnable PS, et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. Science **326,** 1112-1115.

Schuster-Bockler B, and Bateman A (2005). Visualizing profile-profile alignment: pairwise HMM logos. Bioinformatics **21,** 2912-2913.

Seitz O (2000). Glycopeptide synthesis and the effects of glycosylation on protein structure and activity. Chembiochem : a European journal of chemical biology **1,** 214-246.

Shinohara H, and Matsubayashi Y (2013). Chemical Synthesis of Arabidopsis CLV3 Glycopeptide Reveals the Impact of Hydroxyproline Arabinosylation on Peptide Conformation and Activity. Plant Cell Physiol **54,** 369-374.

Silverstein KAT, Moskal WA Jr, Wu HC, Underwood BA, Graham MA, Town CD, VandenBosch KA (2007) Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. Plant J **51**: 262-280

Smyth GK (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology **3,** Article 3.

Somssich M, and Simon R (2012). Peptides regulating apical meristem. In Plant signaling peptides, H.R. Irving and C. Gehring, eds. (Springer).

Tuskan GA, et al (2006). The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science **313,** 1596-1604.

Vanneste S, De Rybel B, Beemster GT, Ljung K, De Smet I, Van Isterdael G, Naudts M, Iida R, Gruissem W, Tasaka M, Inze D, Fukaki H, and Beeckman T (2005). Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in Arabidopsis thaliana. Plant Cell **17,** 3035-3050.

Whitford R, Fernandez A, Tejos R, Perez AC, Kleine-Vehn J, Vanneste S, Drozdzecki A, Leitner J, Abas L, Aerts M, Hoogewijs K, Baster P, De Groodt R, Lin YC, Storme V, Van de Peer Y, Beeckman T, Madder A, Devreese B, Luschnig C, Friml J, and Hilson P (2012). GOLVEN Secretory Peptides Regulate Auxin Carrier Turnover during Plant Gravitropic Responses. Dev Cell **22,** 678-685.

Zhang HT, Zhou HJ, Berke L, Heck AJR, Mohammed S, Scheres B, and Menke FLH (2013). Quantitative Phosphoproteomics after Auxin-stimulated Lateral Root Induction Identifies an SNX1 Protein Phosphorylation Site Required for Growth. Mol Cell Proteomics **12,** 1158-1169.

Tables

Table 1. Plant secretory peptides and their roles in root development

Peptide family	Functions	Family ID	Arabidopsis*	Rice***	Poplar	Maize	Grapevine	References
CLE	RAM maintenance, vascular development	f5, f9	30 (32)	45;38	48	44	1	Stahl et al., 2009; Kiyohara and Sawa, 2012
IDA	Lateral root emergence	f7	7 (6)	4;3	12	5	0	Kumpf et al., 2013
PSK	QC cell division	f53	8 (6)	6;6	10	9	6	Heyman et al., 2013
PSY	Cell elongation	f74, f4335	16 (3)	13;10	10	12	3	Amano et al. (2007)
RALF	Growth, rhizosphere acidification	f19, f839, f4248	40 (34)	19;24	23	30	9	Srivastava et al., 2009
CEP1	Growth and branching	f195	6 (15**)	4;4	5	4	1	Delay et al., 2013b; Roberts et al., 2013
CEP2		f35	4	5;9	3	7	0	
GLV/RGF/CLEL	Lateral root formation, RAM maintenance, hair growth, gravitropism	f2	12	8;11	12	17	0	Matsuzaki et al., 2010; Whitford et al., 2012; Fernandez et al., 2013a
GASA	GA signaling, cell division (?)	f290	18 (15)	10;15	19	15	9	Roxrud et al. (2007)
New family 1	LR development	f31	4	6;2	5	5	0	This study
New family 2	LR development	f919	2	3;2	1	9	0	This study
New family 3	LR development	f1528	3	1;1	8	1	0	This study

* Number of previously described Arabidopsis peptides assembled in this study in the corresponding families. Peptides of the same family annotated in the Arabidopsis genome annotation TAIR10 are listed in parentheses.

** Four new CEP genes identified in the listed papers were not annotated in TAIR10. CEPs have been classified in a single family but our study separates them into two.

*** Two rice genome annotations provided complementary predicted SSPs: left numbers from RAP-DB, right from MSU6.1. Update of the table

Table 2. Specific spatiotemporal expression of uncharacterized SSP genes during lateral root initiation

Peptide name[a]	AGI ID	ATH1 probeset	SLR-dependent auxin pathway		Auxin & naxillin pathway		Xylem pole pericycle	Pericycle diff. expr.	Radial layers	DR5 oscill.[d]
			Auxin induction[a]	SLR dependence	Auxin indution[b]	Naxillin induction[b]	Auxin induction[b]			
1-1	AT3G06090	256391_at								
1-2	AT4G37295	253047_at	0 to 2 hours	Yes	0 to 2 hours	0 to 6 hours	0 to 2 hours		Primordium	
1-3	AT4G28460	253796_at						PPP[c]		P2
2-2	AT4G34600	253246_at	2 to 6 hours	Yes	0 to 2 hours		2 to 6 hours			
3-1	AT1G49800	259809_at								P5
3-2	AT2G23270	245082_at			0 to 2 hours					
3-3	AT4G37290	253044_at	0 to 2 hours	Yes	0 to 2 hours	0 to 2 hours	0 to 2 hours			

[a] The first number indicates the new family number. Correspondence with family names in Table 1 and Figure 3: 1, f31; 2, f919; 3, f1528.

[b] Indicates time after treatment: between 0 and 2 hours (early transition), 2 and 6 hours (late transition) or 0 and 6 hours (slow transition).

[c] PPP, phloem pole pericycle layer.

[d] Px indicates a cluster in phase with DR5 oscillations.

Supporting Information

Supplemental Table S1. SSP genes collected for testing *de novo* secretory peptide detection strategy.

See Excel file

Supplemental Table S2. Specific expression patterns of known SSP genes during LR formation.

See Excel file

Key to Supplemental Table S2. Specific expression patterns of known SSP genes during LR formation

For each additionally discovered SSP, the following are indicated: the corresponding AGI identifier, the Affymetrix probe set, and the number of datasets in which the gene is significantly differentially expressed in relation to the LR initiation with a fold change ≥ 1.5 and *P*-value ≤ 0.01. “nd”, no data; “0-2”, “2-6” and “0-6”, significant transition between 0 and 2 h (early transition), 2 and 6 h (late transition), or 0 and 6 h after treatment (slow transition), respectively; PPP, phloem pole pericycle layer; XPP, xylem pole pericycle layer. Information about the other layers (APL, S32, SUC2, SCR5, S18, COBL9, AGL42) are as in (Brady et al., 2007). Px and Ax indicate clusters in phase or in antiphase with DR5 oscillations.

Supplemental Table S3. Synthetic peptide sequences tested for effect on root growth and development

Group	Peptide #	AGI ID	Specific LR pattern[a]	Amino acid sequence
1	1-1	AT3G06090	no	WSGRKLASGPSRRGCGH
1	1-2	AT4G37295	yes	MHTMASGPSRRGAGH
1	1-3	AT4G28460	no	LASGPSPRGRGH
1	1-4	AT5G43066	nd	MRILASGPNKRGRGH
2	2-1	AT2G16385	nd	DYGNNSPSPRLERPPFKLIPN
2	2-2	AT4G34600	yes	DYGHSSPKPKLVRPPFKLIPN
3	3-1-1	AT1G49800	no	LATVKQSGPSPGVGHH
3	3-1-2	AT1G49800	no	DSGPSPGVGH
3	3-2-1	AT2G23270	no	LGAIKESGPSSGGEGH
3	3-2-2	AT2G23270	no	HSGPSTSGPGH
3	3-3-1	AT4G37290	yes	LGSIKDSGPSPGEGH
3	3-3-2	AT4G37290	yes	HSGPSPSGPGH

[a] As determined in Table 2.

Supplemental Table S1. SSP genes collected as a benchmark set for *de novo* secretory peptide detection algorithms.

Propeptide	LOCUS	Name	References
CEP	AT1G16950	CEP13	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT1G29290	CEP14	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT1G31670b	CEP12	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT1G47485	CEP1	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT1G59835	CEP2	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT2G23440	CEP3	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT2G35612	CEP4	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT2G40530	CEP15	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT3G50610	CEP9	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT5G66815	CEP5	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	AT5G66816	CEP6	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	Between AT1G36040-AT1G36050	CEP10	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	Between AT2G23440-AT2G23450	CEP11	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CEP	Between AT5G66817-	CEP7	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331

	AT5G66820		
CEP	Between AT5G66817-AT5G66820	CEP8	J. Exp. Bot. (2013) 64 (17): 5371-5381. doi: 10.1093/jxb/ert331
CLE	AT1G05065	CLE20	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G06225	CLE3	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G25425	CLE43	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G26600	CLE9	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G49005	CLE11	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G63245	CLE14	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G66145	CLE18	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G67775	CLE8	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G68795	CLE12	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G69320	CLE10	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G69588	CLE45	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G69970	CLE26	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G70895	CLE17	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G73165	CLE1	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT1G73965	CLE13	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5

CLE	AT2G01505	CLE16	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT2G27250	CLV3.1	Science. 1999 Mar 19;283(5409):1911-4
CLE	AT2G31081	CLE4	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT2G31082	CLE7	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT2G31083	CLE5	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT2G31085	CLE6	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT2G34925	CLE42	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT3G24225	CLE19	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT3G24770	CLE41	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT3G25905	CLE27	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT3G28455	CLE25	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT4G13195	CLE44	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT4G18510	CLE2	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT5G12235	CLE22	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT5G12990	CLE40	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT5G59305	CLE46	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
CLE	AT5G64800	CLE21	Cell. Mol. Life Sci. DOI 10.1007/s00018-007-7411-5
DVL/RTFL	AT1G07490	DEVIL 9; DVL9; ROTUNDIFOLIA LIKE	TAIR10

		3; RTFL3	
DVL/RTFL	AT1G13245	DEVIL 4; DVL4; ROTUNDIFOLIA LIKE 17; RTFL17	TAIR10
DVL/RTFL	AT1G17235	ROTUNDIFOLIA LIKE 11; RTFL11	TAIR10
DVL/RTFL	AT1G53708	ROTUNDIFOLIA LIKE 9; RTFL9	TAIR10
DVL/RTFL	AT1G64585	DEVIL 12; DVL12; ROTUNDIFOLIA LIKE 22; RTFL22	TAIR10
DVL/RTFL	AT1G67265	DEVIL 3; DVL3; ROTUNDIFOLIA LIKE 21; RTFL21	TAIR10
DVL/RTFL	AT1G68825	DEVIL 5; DVL5; ROTUNDIFOLIA LIKE 15; RTFL15	TAIR10
DVL/RTFL	AT2G29125	DEVIL 13; DVL13; ROTUNDIFOLIA LIKE 2; RTFL2	TAIR10
DVL/RTFL	AT2G36985	DEVIL 16; DVL16; ROT4; ROTUNDIFOLIA4	TAIR10
DVL/RTFL	AT2G39705	DEVIL 11; DVL11; ROTUNDIFOLIA LIKE 8; RTFL8	TAIR10
DVL/RTFL	AT3G02493	DEVIL 2; DVL2; ROTUNDIFOLIA LIKE 19; RTFL19	TAIR10
DVL/RTFL	AT3G14362	DEVIL 19; DVL19; ROTUNDIFOLIA LIKE 10; RTFL10	TAIR10
DVL/RTFL	AT3G18518	DEVIL 7; DVL7; ROTUNDIFOLIA LIKE 20; RTFL20	TAIR10
DVL/RTFL	AT3G23635	ROTUNDIFOLIA LIKE 13; RTFL13	TAIR10

DVL/RTFL	AT3G23637	DEVIL 21; DVL21	TAIR10
DVL/RTFL	AT3G25717	DEVIL 6; DVL6; ROTUNDIFOLIA LIKE 16; RTFL16	TAIR10
DVL/RTFL	AT3G46613	DEVIL 15; DVL15; ROTUNDIFOLIA LIKE 4; RTFL4	TAIR10
DVL/RTFL	AT3G53232	DEVIL 20; DVL20; ROTUNDIFOLIA LIKE 1; RTFL1	TAIR10
DVL/RTFL	AT3G55515	DEVIL 8; DVL8; ROTUNDIFOLIA LIKE 7; RTFL7	TAIR10
DVL/RTFL	AT3G63088	DEVIL 14; DVL14; ROTUNDIFOLIA LIKE 14; RTFL14	TAIR10
DVL/RTFL	AT4G13395	DEVIL 10; DVL10; ROTUNDIFOLIA LIKE 12; RTFL12	TAIR10
DVL/RTFL	AT4G35783	DEVIL 17; DVL17; ROTUNDIFOLIA LIKE 6; RTFL6	TAIR10
DVL/RTFL	AT5G16023	DEVIL 1; DVL1; ROTUNDIFOLIA LIKE 18; RTFL18	TAIR10
DVL/RTFL	AT5G59510	DEVIL 18; DVL18; ROTUNDIFOLIA LIKE 5; RTFL5	TAIR10
EPF	AB4499313	EPFL7	new
EPF	AB499312	EPFL3	new
EPF	AT1G34245	EPF2; EPIDERMAL PATTERNING FACTOR 2	TAIR10
EPF	AT1G71866		new
EPF	AT1G80133	EPFL8	new
EPF	AT2G20875	EPF1; EPIDERMAL PATTERNING FACTOR	TAIR10

		1	
EPF	AT2G30370	CHAL; CHALLAH; EPF1-LIKE 6; EPFL6	TAIR10
EPF	AT3G13898		new
EPF	AT3G22820	CHALLAH-LIKE 1; CLL1; EPFL5; EPIDERMAL PATTERNING FACTOR LIKE 5	TAIR10
EPF	AT4G12970	EPFL9; STOMAGEN	TAIR10
EPF	AT4G14723	CHALLAH-LIKE 2; CLL2; EPFL4; EPIDERMAL PATTERNING FACTOR LIKE 4	TAIR10
EPF	AT4G37810	EPFL2	new
EPF	AT5G10310	EPFL1	new
GASA	AT1G10588	GASA-related	Plant Cell Physiol. 48(3):471-83
GASA	AT1G22690	GASA09	Plant Cell Physiol. 48(3):471-83
GASA	AT1G74670	GASA06	Plant Cell Physiol. 48(3):471-83
GASA	AT1G75750	GASA01	Plant Cell Physiol. 48(3):471-83
GASA	AT2G14900	GASA07	Plant Cell Physiol. 48(3):471-83
GASA	AT2G18420	GASA11	Plant Cell Physiol. 48(3):471-83
GASA	AT2G30810	GASA12	Plant Cell Physiol. 48(3):471-83
GASA	AT2G39540	GASA08	Plant Cell Physiol. 48(3):471-83
GASA	AT3G02885	GASA05	Plant Cell Physiol. 48(3):471-83
GASA	AT3G10185	GASA13	Plant Cell Physiol. 48(3):471-83
GASA	AT4G09600	GASA03	Plant Cell Physiol. 48(3):471-83
GASA	AT4G09610	GASA02	Plant Cell Physiol. 48(3):471-83
GASA	AT5G14920	GASA14	Plant Cell Physiol. 48(3):471-83
GASA	AT5G15230	GASA04	Plant Cell Physiol. 48(3):471-83

GASA	AT5G59845	GASA10	Plant Cell Physiol. 48(3):471-83
GLV	AT1G13620	GLV5/RGF2/CLEL1	TAIR10
GLV	AT2G03830	GLV6/RGF8/CLEL2	TAIR10
GLV	AT2G04025	GLV7/RGF3/CLEL3	TAIR10
GLV	AT3G02240	GLV4/RGF7/CLEL4	TAIR10
GLV	AT3G02242	GLV8/CLEL5	TAIR10
GLV	AT3G30350	GLV3/RGF4	TAIR10
GLV	AT4G16515	GLV1/RGF6/CLEL6	TAIR10
GLV	AT5G15725	GLV9	TAIR10
GLV	AT5G51451	GLV10/RGF5/CLEL7	TAIR10
GLV	AT5G60810	GLV11/RGF1/CLEL8	TAIR10
GLV	AT5G64770	GLV2/RGF9/CLEL9	TAIR10
IDL	AT1G76952	IDL5	TAIR10
IDL	AT3G18715	IDL4	TAIR10
IDL	AT3G25655	IDL1	TAIR10
IDL	AT4G18335		Plant Cell. 2003 Oct;15(10):2296-307
IDL	AT5G09805	IDL3	TAIR10
IDL	AT5G64667	IDL2	TAIR10
PNP	AT2G18660	ATPNP-A	TAIR10
PNP	AT4G30380	ATPNP-B	TAIR10
POLARIS	AT4G39403	POLARIS	TAIR10
PROPEP	AT1G17750	ATPEPR2	TAIR10
PROPEP	AT1G73080	ATPEPR1	TAIR10
PROPEP	AT2G22000	PROPER6	TAIR10
PROPEP	AT5G09978	PROPER7	TAIR10
PROPEP	AT5G09980	PROPER4	TAIR10
PROPEP	AT5G09990	PROPER5	TAIR10

PROPEP	AT5G64905	PROPER3	TAIR10
PSK	AT1G13590	PSK1	TAIR10
PSK	AT2G22860	PSK2	TAIR10
PSK	AT2G22942	PSK-related	TAIR8
PSK	AT3G44735	PSK3	TAIR10
PSK	AT3G49780	PSK4	TAIR10
PSK	AT4G37720	PSK6	TAIR10
PSK	AT5G65870	PSK5	TAIR10
PSY	AT1G07175	PSY1	TAIR8
PSY	AT1G74458	PSY1	TAIR8
PSY	AT2G29995	PSY1	TAIR8
PSY	AT3G47295	PSY1	TAIR8
PSY	AT3G47510	PSY1	TAIR8
PSY	AT3G49270	PSY1	TAIR8
PSY	AT3G49300	PSY1	TAIR8
PSY	AT3G49305	PSY1	TAIR8
PSY	AT3G49307	PSY1	TAIR8
PSY	AT5G53486	PSY1	TAIR8
PSY	AT5G58650	PSY1	TAIR10
RALFL	AT1G02900	RALFL1	TAIR10
RALFL	AT1G23145	RALFL2	TAIR10
RALFL	AT1G23147	RALFL3	TAIR10
RALFL	AT1G28270	RALFL4	TAIR10
RALFL	AT1G35467	RALFL5	TAIR10
RALFL	AT1G60625	RALFL6	TAIR10
RALFL	AT1G60815	RALFL7	TAIR10
RALFL	AT1G60835	pseudogene	TAIR10

RALFL	AT1G60913		TAIR10
RALFL	AT1G61563	RALFL8	TAIR10
RALFL	AT1G61566	RALFL9	TAIR10
RALFL	AT2G19020	RALFL10	TAIR10
RALFL	AT2G19030	RALFL11	TAIR10
RALFL	AT2G19040	RALFL12	TAIR10
RALFL	AT2G19045	RALFL13	TAIR10
RALFL	AT2G20660	RALFL14	TAIR10
RALFL	AT2G22055	RALFL15	TAIR10
RALFL	AT2G32785		TAIR10
RALFL	AT2G32788		TAIR10
RALFL	AT2G32835	RALFL16	TAIR10
RALFL	AT2G32885		TAIR10
RALFL	AT2G32890	RALFL17	TAIR10
RALFL	AT2G33130	RALF18	TAIR10
RALFL	AT2G33775	RALFL19	TAIR10
RALFL	AT2G34825	RALFL20	TAIR10
RALFL	AT3G04735	RALFL21	TAIR10
RALFL	AT3G05490	RALFL22	TAIR10
RALFL	AT3G16570	RALF23	TAIR10
RALFL	AT3G23805	RALFL24	TAIR10
RALFL	AT3G25165	RALFL25	TAIR10
RALFL	AT3G25170	RALFL26	TAIR10
RALFL	AT3G29780	RALFL27	TAIR10
RALFL	AT4G11510	RALFL28	TAIR10
RALFL	AT4G11653	RALFL29	TAIR10
RALFL	AT4G13075	RALFL30	TAIR10

RALFL	AT4G13950	RALFL31	TAIR10
RALFL	AT4G14010	RALFL32	TAIR10
RALFL	AT4G14020		TAIR10
RALFL	AT4G15800	RALFL33	TAIR10
RALFL	AT5G67070	RALFL34	TAIR10
SCRL	AT1G08695	SCR-LIKE 3; SCRL3	TAIR10
SCRL	AT1G14182	SCR-LIKE 28; SCRL28	TAIR10
SCRL	AT1G60983	SCR-LIKE 8; SCRL8	TAIR10
SCRL	AT1G60985	SCR-LIKE 6; SCRL6	TAIR10
SCRL	AT1G60986	SCR-LIKE 4; SCRL4	TAIR10
SCRL	AT1G60987	SCR-LIKE 5; SCRL5	TAIR10
SCRL	AT1G60989	SCR-LIKE 7; SCRL7	TAIR10
SCRL	AT1G65113	SCR-LIKE 2; SCRL2	TAIR10
SCRL	AT2G05117	SCR-LIKE 9; SCRL9	TAIR10
SCRL	AT2G05335	SCR-LIKE 15; SCRL15	TAIR10
SCRL	AT2G06983	SCR-LIKE 16; SCRL16	TAIR10
SCRL	AT2G14282	SCR-LIKE 18; SCRL18	TAIR10
SCRL	AT2G25685	SCR-LIKE 17; SCRL17	TAIR10
SCRL	AT3G23715	SCR-LIKE 13; SCRL13	TAIR10
SCRL	AT3G23727	SCR-LIKE 12; SCRL12	TAIR10
SCRL	AT3G27503	SCR-LIKE 19; SCRL19	TAIR10
SCRL	AT4G10115	SCR-LIKE 20; SCRL20	TAIR10
SCRL	AT4G10457	SCR-LIKE 1; SCRL1	TAIR10
SCRL	AT4G10767	SCR-LIKE 21; SCRL21	TAIR10
SCRL	AT4G14785	SCR-LIKE 23; SCRL23	TAIR10
SCRL	AT4G15733	SCR-LIKE 11; SCRL11	TAIR10
SCRL	AT4G15735	SCR-LIKE 10; SCRL10	TAIR10

SCRL	AT4G22105	SCR-LIKE 26; SCRL26	TAIR10
SCRL	AT4G22115	SCR-LIKE 14; SCRL14	TAIR10
SCRL	AT4G32714	SCR-LIKE 25; SCRL25	TAIR10
SCRL	AT4G32717	SCR-LIKE 24; SCRL24	TAIR10
SCRL	AT4G33465	SCR-LIKE 22; SCRL22	TAIR10
SCRL	AT5G45875	SCR-LIKE 27; SCRL27	TAIR10
TPD	AT1G05835	TPD1	TAIR8
TPD	AT1G32583	TPD1	TAIR8
TPD	AT4G24972	TPD1	TAIR10
TPD	AT4G32090	TPD1	TAIR8
TPD	AT4G32100	TPD1	TAIR8
TPD	AT4G32105	TPD1	TAIR8
TPD	AT4G32110	TPD1	TAIR8

Supplemental Table S2. Specific expression patterns of known SSP genes during LR formation.

Gene description				Lateral Root Initiation Transcriptomics scoring															
Peptide family	Gene Name	AGI identifier	Affymetrix	Hit count	SLR dependant auxin path			Auxin & Naxillin pathwa			Xylem Pole Pericycle		Pericycle		Radial layers		DR5 oscillations	
			probeset	FC		Auxin inducible			Auxin	Naxillin		Auxin inducible	differential		specificity		Cluster	
				1,5			SLR		inducibl	inducible			expression					
				p-Value	*Vanneste*		dependant	*De Rybel*			*De Smet*		*Parizot*		*Brady*		*Moreno-Rissueno*	
				0,01	*et al., 2005*			*et al., 2012*			*et al., 2008*		*et al., 2012*		*et al., 2007*		*et al, 2010*	
CEP	CEP1	AT1G47485	262445_at	0														
CEP	CEP12	AT1G31670	246574_at	1													HIT	P5
CEP	CEP1-like	AT5G66815	247070_at	1						0-2					HIT	XPP		
CEP	CEP1-like	AT2G23440	267133_at	0														
CEP	CEP1-like	AT1G59835		nd														
CEP	CEP1-like	AT2G35612		nd														
CEP	CEP2	AT1G16950	256109_at	0														
CEP	CEP2-like	AT5G05300	250796_at	0														
CEP	CEP2-like	AT2G40530	255824_at	0		2-6			2-6									
CEP	CEP2-like	AT1G29290		nd														
CEP	CEP6	AT5G66816		nd														
CEP	CEP9	AT3G50610	252163_at	0														
CLE	CLE01	AT1G73165		nd														
CLE	CLE02	AT4G18510	254644_at	2		0-2			0-2		HIT	0-2			HIT	PPP		
CLE	CLE03	AT1G06225	260795_at	1													HIT	P5
CLE	CLE04	AT2G31081		nd														
CLE	CLE05	AT2G31083		nd														
CLE	CLE06	AT2G31085	266477_at	0												APL,S32		
CLE	CLE07	AT2G31082		nd														
CLE	CLE08	AT1G67775		nd														
CLE	CLE09	AT1G26600	261012_at	0														
CLE	CLE10	AT1G69320		nd														
CLE	CLE11	AT1G49005		nd														

CLE	CLE12	AT1G68795	260039_at	1				HIT	2-6	0-2									
CLE	CLE13	AT1G73965		nd															
CLE	CLE14	AT1G63245		nd															
CLE	CLE16	AT2G01505		nd															
CLE	CLE17.1	AT1G70895	262261_at	0															
CLE	CLE18	AT1G66145		nd															
CLE	CLE19	AT3G24225		nd															
CLE	CLE20	AT1G05065		nd															
CLE	CLE21	AT5G64800	247254_at	0															
CLE	CLE22	AT5G12235		nd															
CLE	CLE25	AT3G28455		nd															
CLE	CLE26	AT1G69970	264693_at	0					0-2							APL, S32			
CLE	CLE27	AT3G25905	258082_at	0															
CLE	CLE40	AT5G12990	250271_at	0															
CLE	CLE41	AT3G24770	257600_at	1						2-6			HIT	PPP		APL,SUC2,S			
CLE	CLE42	AT2G34925		nd															
CLE	CLE43	AT1G25425		nd															
CLE	CLE44	AT4G13195	254761_at	3		0-2		HIT	0-2	2-6	HIT	0-6					HIT	P1	
CLE	CLE45	AT1G69588		nd															
CLE	CLE46	AT5G59305	247719_at	0															
CLE	CLV3.1	AT2G27250	265624_at	0															
GASA	GASA01	AT1G75750	262947_at	1		2-6			2-6		HIT	0-2							
GASA	GASA02	AT4G09610	255049_at	0															
GASA	GASA03	AT4G09600	255048_at	0															
GASA	GASA04	AT5G15230	250109_at	1									HIT	XPP					
GASA	GASA05	AT3G02885	258618_at	1		0-2			0-2		HIT	0-2							
GASA	GASA06	AT1G74670	260221_at	2	HIT	2-6	SLR dep		0-2		HIT	0-2							
GASA	GASA07	AT2G14900	266613_at	2		0-2		HIT	2-6	0-2	HIT	2-6				SCR5			
GASA	GASA08	AT2G39540	266969_at	0															
GASA	GASA09	AT1G22690	264195_at	1													HIT	P5	

GASA	GASA10	AT5G59845	247657_at	1				HIT	2-6	0-2						S18		
GASA	GASA11	AT2G18420	265331_at	2							HIT	0-6					HIT	P5
GASA	GASA12	AT2G30810	267198_at	0														
GASA	GASA13	AT3G10185		nd														
GASA	GASA14	AT5G14920	246550_at	0														
GASA	GASA-related	AT1G10588		nd														
GLV/RGF/CLEL	GLV01	AT4G16515	245336_at	1													HIT	P5
GLV/RGF/CLEL	GLV02	AT5G64770	247252_at	0														
GLV/RGF/CLEL	GLV03	AT3G30350	256710_at	0														
GLV/RGF/CLEL	GLV04	AT3G02240	259120_at	0		0-2			0-2							COBL9		
GLV/RGF/CLEL	GLV05	AT1G13620		nd														
GLV/RGF/CLEL	GLV06	AT2G03830	263360_at	2		0-2			0-2		HIT	0-2				AGL42	HIT	P4
GLV/RGF/CLEL	GLV07	AT2G04025	263481_at	0														
GLV/RGF/CLEL	GLV08	AT3G02242		nd														
GLV/RGF/CLEL	GLV09	AT5G15725	246530_at	1													HIT	P3
GLV/RGF/CLEL	GLV10	AT5G51451		nd														
GLV/RGF/CLEL	GLV11.1	AT5G60810		nd														
IDL	IDA	AT1G68765	260040_at	1	HIT	2-6	SLR dep		2-6									
IDL	IDA related	AT4G18335		nd														
IDL	IDL1	AT3G25655	256762_at	0						0-2								
IDL	IDL2	AT5G64667		nd														
IDL	IDL3	AT5G09805		nd														
IDL	IDL4	AT3G18715		nd														
IDL	IDL5	AT1G76952		nd														
PNP	PNP-A	AT2G18660	266070_at	1					2-6		HIT	0-2						
PNP	PNP-B?	AT4G30380	253616_at	1													HIT	P5
PROPEP	ATPEPR1	AT1G73080	262360_at	1		2-6			2-6				HIT	PPP				
PROPEP	ATPEPR2	AT1G17750	259400_at	2		0-2			0-2		HIT	2-6					HIT	P2
PROPEP	PROPEP1	AT5G64900	247213_at	3		0-6			2-6		HIT	2-6	HIT	XPP			HIT	P2
PROPEP	PROPEP2	AT5G64890	247205_at	0					2-6									

PROPEP	PROPEP3	AT5G64905	247215_at	2		0-2		HIT	0-2	0-2	HIT	2-6							
PROPEP	PROPEP4	AT5G09980	250455_at	2				HIT	2-6	0-2	HIT	0-2							
PROPEP	PROPEP5	AT5G09990	250485_at	0					0-6										
PROPEP	PROPEP6	AT2G22000	263869_at	0					2-6										
PROPEP	PROPER7	AT5G09978		nd															
PSK	PSK1	AT1G13590	256158_at	0					0-6										
PSK	PSK2	AT2G22860	266799_at	2		2-6			0-2		HIT	0-6	HIT	PPP					
PSK	PSK3	AT3G44735	252624_at	2		2-6		HIT	0-2	0-2			HIT	XPP					
PSK	PSK4	AT3G49780	252234_at	2	HIT	0-2	SLR dep		0-2				HIT	PPP		SCR5			
PSK	PSK5	AT5G65870	247109_at	2	HIT	2-6	SLR dep		2-6				HIT	PPP		COBL9			
PSK	PSK6	AT4G37720		nd															
PSK	PSK-related	AT2G22942		nd															
PSY	PSY related	AT3G49300	252253_at	0															
PSY	PSY related	AT3G49270	252306_at	0															
PSY	PSY related	AT3G47295	252412_at	0															
PSY	PSY related	AT3G47510	252419_at	2	HIT	2-6	SLR dep		2-6				HIT	PPP		SUC2			
PSY	PSY related	AT2G29995	266808_at	0															
PSY	PSY related	AT1G07175		nd															
PSY	PSY related	AT1G74458		nd															
PSY	PSY related	AT3G49305		nd															
PSY	PSY related	AT3G49307		nd															
PSY	PSY related	AT5G53486		nd															
PSY	PSY1	AT5G58650	247793_at	1		0-2			2-6				HIT	PPP					
RALFL	RALFL01	AT1G02900	262131_at	1		0-2			0-2				HIT	PPP					
RALFL	RALFL02	AT1G23145		nd															
RALFL	RALFL03	AT1G23147		nd															
RALFL	RALFL04	AT1G28270	245658_at	0															
RALFL	RALFL05	AT1G35467		nd															
RALFL	RALFL06	AT1G60625		nd															
RALFL	RALFL07	AT1G60815		nd															

RALFL	RALFL08	AT1G61563		nd															
RALFL	RALFL09	AT1G61566		nd															
RALFL	RALFL10	AT2G19020		nd															
RALFL	RALFL11	AT2G19030		nd															
RALFL	RALFL12	AT2G19040		nd															
RALFL	RALFL13	AT2G19045		nd															
RALFL	RALFL14	AT2G20660	263740_at	0															
RALFL	RALFL15	AT2G22055		nd															
RALFL	RALFL16	AT2G32835		nd															
RALFL	RALFL17	AT2G32890	267643_at	0															
RALFL	RALFL18	AT2G33130	245158_at	1														HIT	P5
RALFL	RALFL19	AT2G33775		nd															
RALFL	RALFL20	AT2G34825		nd															
RALFL	RALFL21	AT3G04735		nd															
RALFL	RALFL22	AT3G05490	259106_at	1	HIT	0-2	SLR dep			2-6									
RALFL	RALFL23	AT3G16570	258432_at	2		2-6				2-6	HIT	0-2	HIT	XPP		S18			
RALFL	RALFL24	AT3G23805	257204_at	1					2-6									HIT	A2
RALFL	RALFL25	AT3G25165	257819_at	0															
RALFL	RALFL26	AT3G25170	257821_at	0															
RALFL	RALFL27	AT3G29780	256563_at	0					0-2										
RALFL	RALFL28	AT4G11510	254900_at	1														HIT	P5
RALFL	RALFL29	AT4G11653		nd															
RALFL	RALFL30	AT4G13075		nd															
RALFL	RALFL31	AT4G13950	245310_at	0					2-6										
RALFL	RALFL32	AT4G14010	245386_at	1		2-6			2-6		HIT	2-6							
RALFL	RALFL33	AT4G15800	245334_at	2		0-2		HIT	2-6	2-6								HIT	A1
RALFL	RALFL34	AT5G67070	247037_at	1		2-6			2-6		HIT	2-6							
RALFL	RALF-related	AT4G14020	245385_at	0												APL,SUC2,S			
RALFL	RALF-related	AT5G38980	249482_at	0					0-6										
RALFL	RALF-related	AT1G60913		nd															

RALFL	RALF-related	AT2G32785		nd														
RALFL	RALF-related	AT2G32788		nd														
RALFL	RALF-related	AT2G32885		nd														
RALFL	RALF-related	AT1G60835		nd														
RALFL	RALF-related	AT1G24145		nd														

Chapter 6

Conclusions and Perspectives

Conclusions and Perspectives

Regardless of our limited knowledge on small signaling peptides in the plant field, their importance and requirement for plant growth and development are well recognized. The plant peptide signaling research is currently maturing and steadily illustrating the complexity of the signaling peptide mechanisms that coordinate growth and development. Our study led to the identification of several novel signaling peptides in *Arabidopsis* that are involved in LR development. Moreover, in this study, we focused on *GOLVEN (GLV)*, a novel peptide family that regulates important developmental aspects in *Arabidopsis*. Our research has contributed to the discovery of upstream components of the GLV signaling pathway and to the characterization of the GLV signaling peptide function during root and lateral root (LR) development.

A major obstacle in the study and functional characterization of relevance of signaling peptides for plant growth and development has been the identification of the genes encoding them. The bioinformatic approach described here presents a valuable tool to discover novel small secreted peptides (SPPs). Retrieval of all known signaling peptide families using our algorithm demonstrates its reliability. Successive combination of SSP gene annotation, expression studies and *in vivo* peptide assays provides, as is demonstrated by this thesis, an original approach to identify potential novel signaling peptides. This type of assay is cheap, easy and fast. Moreover, it can be rapidly and regularly updated as information accrues. Applying this combinatory method can lead to the identification of new small secreted peptides involved in specific developmental aspects and can be adapted to a wide range of processes.

Our study led to identification of several novel SSPs potentially involved in LR development. One of which, AT4G34600 showed strong inhibition of LR emergence when exogenously applied. Nothing else is known about this protein. Our primary data suggest a role in LR development but further detailed studies are required to elucidate its function. For instance, determining its expression pattern in different tissue and cell types and during different developmental stages, will give us an idea about the specificity of this peptide's function. The next step will be the phenotypic analysis of gain-of-function and loss-of-function mutants which will be essential to obtain a better insight into its developmental function. Monitoring different stages of LRP development in these lines is crucial to dissect how AT4G34600 regulates LRP growth. Once a potential role during lateral root formation has been validated by the gain- and

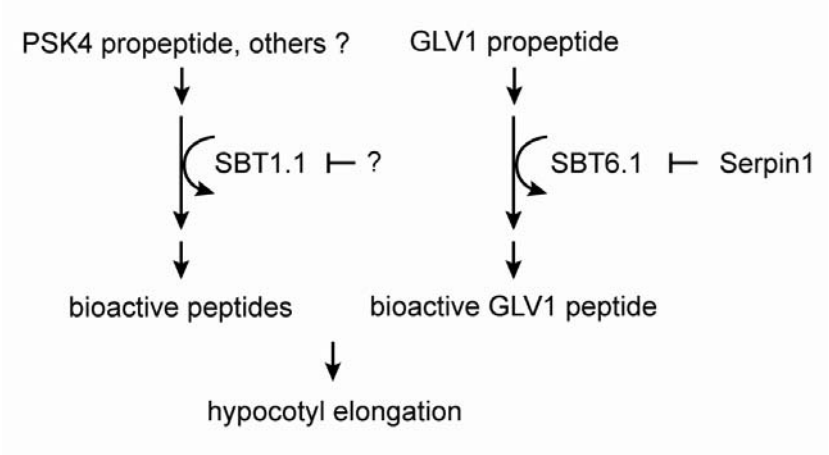
loss-of-function approach, the identification of upstream and downstream components of the AT4G34600 signaling pathway would be our long term plan to follow up on this project.

The above mentioned sequence comparison algorithms led to the identification of a new SSP family named GOLVEN (GLV). GLV proteins are recognized as small signaling peptides and their precursors require proteolytic processing to yield the mature peptide. As another part of our work, we looked for enzyme(s) responsible for the proteolytic cleavage of GLV peptides. To this end, we started a suppressor screen based on *GLV1*-overexpression root phenotypes. Our initial screen and the following analysis of related hypocotyl growth phenotypes revealed that the genes coding for the subtilases SBT6.1 and SBT6.2 are essential for maturation and activation of the GLV1 peptide. Interestingly, further studies revealed that SBT6.1 is also under the control of the Serpin1 protease inhibitor.

In vitro protease assays revealed that the plant-purified SBT6.1 enzyme cleaves the GLV1 precursor peptides at sites reminiscent of the canonical recognition sequences for subtilases, RXXL and RXLX. As 10 out of 11 GLV precursors carry at least one of these sites, SBT6.1 may cleave multiple members of the GLV family. To test this hypothesis, monitoring the *in vitro* processing of other GLV synthetic peptides by SBT6.1 will be essential. Besides, our biochemical analysis revealed that SBT6.1 does not cleave the GLV1 precursor sequence immediately upstream of the mature GLV1 peptide. Therefore, additional enzymes, such as SBT6.2, are possibly required to produce the mature bioactive GLV1 peptide. Remarkably, the mammalian counterpart of SBT6.2, TPP II is a known exopeptidase, which removes tripeptides from the amino-termini of larger peptides. Possibly, SBT6.2 might show a similar function in plants and cleaves the extra residues in GLV peptides. However, this needs to be further investigated.

In contrast to the relatively broad expression patterns of the *SBT6.1* and *Serpin1* genes in *Arabidopsis*, the expression patterns of the respective *GLV* genes are much more confined to certain tissues or cell types. We hypothesize that SBT6.1 and Serpin1 might be components of a basic GLV processing machinery available in most of the tissues. Therefore, they may be involved in the co-regulation of *GLV* functions in various tissues, including, but not exclusively, the hypocotyl.

Our observations confirm that *GLV1* and *GLV2* play a positive role in the cell expansion regulation in agreement with their requirement for the gravitropic response of reoriented hypocotyls, a process during which the differential regulation of cell elongation is crucial. Although significant, the differences we observed in hypocotyl lengths are limited to 10-20% in gain- or loss-of-function when compared to the wild type and it is obvious that GLVs are not the sole master regulators of hypocotyl elongation. In fact, other secreted peptides have been shown to promote hypocotyl cell expansion, including PHYTOSULFOKINEα (PSKα) and PLANT PEPTIDE CONTAINING SULFATED TYROSINE 1 (PSY1). Noticeably, mature GLV, PSK and PSY peptides all carry a sulfated tyrosine residue that is important for bioactivity and they may also share other processing enzymes, including subtilases. Cell elongation is a process known to be dependent on plant hormones such as auxin, ethylene and gibberellins. Yet, a possible crosstalk between the peptide signaling pathways that drive cell expansion remains to be elucidated, as well as the connection with hormonal growth control.



Since GLV peptides have been suggested to be involved in numerous developmental features, including root and LR development, we focused also on the characterization of the GLV signaling peptide family during primary root and LR development. Morphological analysis revealed that overexpression of *GLV* genes can perturb normal growth and emergence of the LRP. The spatiotemporal regulation of cell division has a great impact on the fate of a primordium and our data revealed that GLV peptides might play a role in the coordination of cell division at the early stages of LRP development. One of the possible hypotheses is that GLV peptides might act as an upstream component of some of the cell cycle genes, which, if perturbed,

can deregulate the precision of cell divisions during LRP development. It would also be interesting to address genetic interaction between *GLV* and cell cycle genes.

Surprisingly our study showed that loss-of-function of some *GLV* genes can affect LRP development at stages in which the corresponding gene is not expressed. For example loss-of-function of *GLV3* resulted in increased number of stage II LRP whereas *GLV3* is not expressed yet at that stage. This observation might hint at a possible indirect effect of these genes on LRP development. *GLV3* is also expressed in the root apical meristem (RAM) and it was shown that altered GLV3 levels affect auxin transport in the RAM. Therefore GLV3 silencing might indirectly affect LRP formation by altering auxin fluxes in the RAM.

LRs initiate when pericycle cells accumulate auxin, thereby reaching a founder cell status that will trigger asymmetric cell divisions to form a new LRP. In the RAM, GLV peptides regulate auxin fluxes at two stages. In gravistimulated roots, auxin accumulation is coordinated by GLV3 via regulation of PIN2 cellular trafficking and GLV11 defines *PLETHORA1* (*PLT1)* and *PLT2* levels. Recently, it has been demonstrated that both, *PLT* and *PIN-FORMED3* (*PIN3*) play an important role during LR initiation. One of the *GLV* genes expressed early during LR formation is *GLV10* and its overexpression caused severe defects in LRP development, resulting in misshapen primordia. We hypothesize that a similar type of regulation of PIN3 and/or PLT3 proteins could be active during lateral root formation similar to GLV3 and GLV11 function in the RAM. Another member of the PLT family, PLT3 could be regulated by GLV10. Further study is required to detect the auxin maxima in *GLV10* gain-of-function and loss-of-function lines to investigate the likely *GLV10*-dependent regulation of local auxin accumulation during LRP formation.

GLV10 is also expressed in the RAM, therefore a possible function is expected in RAM regulation as well. It has been shown that *GLV5, GLV7* and *GLV11* redundantly control the expression pattern of *PLT1* and *PLT2*. Our data showed that in addition *GLV10* might also contribute in the same way to RAM maintenance, but detailed analysis is needed to further study this hypothesis and to better dissect how *GLV10* regulates RAM size and LR development.

How do the SSPs control different developmental processes? How are different members of the same SSP family involved in different and sometimes even antagonistic functions? What

are the upstream and downstream players of these signaling peptides pathways? How many unidentified SSPs remain yet to be discovered and characterized? And finally, how the knowledge in SSP field can be applied to improve crops? To answer these and other remaining questions, a wide variety of methods and approaches such as genetic screens, transcriptomics, proteomics and cell biology approches will be required. The future research in the coming years should yield more findings that will enable us to elucidate more pieces of the SSPs puzzle and will allow us to draw a more complete picture.

Acknowledgement

First of all, I would like to express my sincere gratitude to my promoters ***Prof. Tom Beeckman*** and ***Dr. Pierre Hilson*** for the continuous support of my Ph.D and research, for their patience, motivation and immense knowledge. ***Tom***, you are someone one will never forget once they meet you. All that I have learned from you is invaluable. Not only I did discover how to become a better scientist, but you've also shown me how to be a better human. ***Pierre***, it is difficult to express my gratefulness for your support during my PhD. This study would not have been possible without your aid and guidance.

In addition to my brilliant promoters, I would like to thank the jury members for their much-appreciated encouragement and insightful comments which improved my thesis.

Ana, I am deeply grateful for all I have learned from you and for your continuous help and support at all stages of this thesis. Because of your devotion and selflessness, you have helped make my thesis a reality, and I truly appreciate your guidance and supervision.

Martine, I would like to express my deep gratitude and respect to you for all your valuable help and for your kindness.

I would like to thank ***Kurt***, ***Annemieke*** and ***Frank*** for their contribution to this thesis.

Living in Ghent provided me with a unique opportunity to meet amazing people from all over the world. ***Ruth, Aga*** *(Agnieszka),* ***Kun***, ***Olfa***, ***Reem***, ***Karl***, ***Youn-Jeong***, ***Xiaohuan***, ***Kristof***, ***Litsa*** *and* ***Anas***, you have my many thanks for all the moments we have shared, and I will always remember the good times I spent in your wonderful presence.

Barbara, ***Beata***, ***Chen***, ***Davy***, ***Dominique***, ***Dora***, ***Franky***, ***Gieljan***, ***Hans***, ***Hanne***, ***Ianto***, ***Long***, ***Marlies***, ***Maria***, ***Michiel***, ***Nino***, ***Rebecca***, ***Riet***, ***Rudy***, ***Wilson***, ***Xuan Wei***, you were all the best lab mates imaginable. Thank you for the stimulating discussions, your valuable help and all the fun we had in the last years.

I knew in advance this would be a long acknowledgment as I have been so blessed to be surrounded by amazing personalities! I express my sincerest apologies that I could not mention each one of them.

To my dear friends: ***Morvarid***, ***Sara***, ***Mansour***, ***Pooneh***, ***Saman***, ***Hadi***, ***Azadeh***, ***Araz***, ***Mahsa***, ***Soren***, ***Samira***, ***Behnaz***, ***Nafiseh***, ***Amir***, ***Niloufar***, ***Maryam***, ***Nima***, ***Zeynab***, ***Leila***, ***Sepehr***, ***Behrouz***, ***Farzaneh***, ***Setareh*** and ***Ehsan***. We had an awesome time together during these years. Thanks for your wonderful friendship and support, and thanks for sharing all those nice moments with me.

Andrzej and ***Boris*** thank you for your friendship, support and being great and reliable people whom I could always talk to about life's difficulties and moments of excitement.

Mehdi, even in Iran, your kind caring attitude made me always feel happy. Thank you for refreshing my thoughts with your amazing poems. I cannot wait to read your new book.

My lovely friend ***Zahra***, my friendship with you is a never-ending journey.

Homeyra, ***Kazem***, ***Alma***, ***Amir-Ali*** and ***Amir-Masoud***, your caring attitude made me feel at home. I was lucky to have your support in Belgium.

Farnoosh, ***Parisa***, ***Dirk*** and ***Kartik***, you are some of the best people I have had the pleasure of meeting! You all are caring, kind, generous and the embodiment of goodness. I am so grateful to know each of you!

Bahareh and ***Hamid***, you are both extraordinarily special to me and made a great impact in my life. Next to my family, you are the dearest souls. Your love and friendship means more to me than you will ever know.

My greatest debt I owe to my amazing family, ***Samarrokh***, ***Bohloul***, ***Radefeh***, ***Fatemeh***, ***Mehdi***, ***Arezoo***, ***Azam***, ***Parisa***, ***Behdokht***, ***Norouz***, ***Shahrouz***, ***Seyyed Javad***, ***Milad***, ***Mohammad***, ***Anna (Zahra)*** and ***Amir***. I am blessed to have such a wonderfully supportive family! Thank you for always believing in me and for your constant and unconditional love. Without your help, I could not have made it.

Thank you!

Sarieh Ghorbani

March 2014

سپاس

در ابتدا، می خواهم صمیمانه از جناب پروفسور **تام یکمان** و جناب دکتر **پیر هیلسون** ، استادهای راهنمایم، برای حمایت همیشگی و همه جانبه‌شان از پژوهش دکتریم تشکر کنم. عزیزانی که با صبر و شکیبایی بی‌اندازه و وجودی سرشار از شور و امید، انگیزه و شوق پویایی و دانستن را در من دمیدند و با نهایت سخاوت و صداقت، دانش بیکرانشان را در اختیارم نهادند. **تام** عزیز، یاد انسانی همانند شما، هرگز از خاطر من فراموش نخواهد شد. احساس می‌کنم، دانشجوی استاد فرهیخته ای چون شما بودن و در کنار شما کار کردن، نه تنها در زمینه علمی برای من روشنگر و راهگشا بوده است، بلکه همواره برایم الهام بخش پیمودن مسیر کمال انسانی خواهد بود. **پیر**، این پایان نامه و این تحقیق بدون راهنمایی و پشتیبانی بی‌شائبه شما هرگز به ثمر نمی نشست.

در کنار سپاسگزاری از استاد های فرهیخته ام، لازم میدانم از تمام اعضای کمیته ی داوران نیز بخاطر همراهی‌ها و نظرات روشنگرشان در جهت ارتقای علمی این پایان نامه تشکر کنم.

آنا فرناندز عزیز، از صمیم قلب برای تمام چیزها ایکه به من آموختی سپاسگزارم. دانش گسترده و راهنمایی تو باعث بهتر شدن و به انجام رسیدن این پژوهش گردید. قدردانی صمیمانه من را بابت تمام کمک هایت پذیرا باش.

مارتینه ی مهربان، احترام و حقشناسی عمیق من را برای تمام لکهای بی ادعای خود در ویرایش این پایان نامه بپذیر.

مایلم در اینجا از **کورت**، **آنه میکه** و **فرانک** بابت همکاری پرباری که در این پژوهش داشتیم تشکر کنم.

زندگی و تحصیل در گنت این فرصت استثنایی رو برای من ایجاد کرد که با انسانهای ارزشمندی از سراسر جهان آشنا بشم. **روت**، **اگا** (اگنیسکا)، **کان**، **الفا**، **ریم**، **کارل**، **یانجونگ**، **شواوان**، **کریستوف**، **لیتسا** و **اناس** از همه شما به خاطر دوستیتان و لحظه‌های خوبی که با هم داشتیم ممنونم.

باربارا، **به آتا**، **چن**، **دیوی**، **دومینیک**، **دورا**، **فرانکی**، **گیلیان**، **هانس**، **هانه**، **ایاتو**، **لانگ**، **مارلیس**، **ماریا**، **میخیل**، **نینو**، **ریت**، **ویلسون** و **شوانگوی**، همه شما بهترین همکارهایی هستین که می‌توان تصور کرد، به خاطر تمام کمک هایتان و لحظات شادی که با هم داشتیم از شما سپاسگزارم.

از ابتدا میدانستم که نوشتن این سپاس نامه کار راحتی نخواهد بود! زندگی من همیشه به‌طور خارق‌العاده‌ای از دوستانی یک‌رنگ و صمیمی سرشار بوده است. دوستانی که سرمایه‌های واقعی زندگی من بوده و هستند و من هیچگاه نمی‌توانم آنطور که آرزو دارم از آنها قدردانی کنم. و البته مجال نیست که به اسامی همه آنها اشاره نمایم. اینجا به اسامی تعدادی از آنها اشاره می‌کنم و به بقیه دوستان عزیزم درود می‌فرستم.

دوستهای خوبم، **مروارید**، **هادی**، **آزاده**، **سارا**، **منصور**، **پونه**، **سامان**، **آراز**، **مهسا**، **سورن**، **سمیرا**، **بهناز**، **نفیسه**، **امیر**، **نیلوفر**، **مریم**، **نیما**، **زینب**، **لیلا**، **سپهر**، **بهروز**، **فرزانه**، **ستاره** و **احسان** برای دوستی صمیمانه و حمایت بی‌دریغتان و تمام لحظه‌های خوبی که با هم داشتیم تشکر می‌کنم.

آندره ی و **بوریس** مهربان، دوستی و مهرشما برایم بی نهایت ارزشمند است. ممنونم که همیشه باصبوری و مهربانی به درد دلهای من گوش کردید و همیشه پشتیبان من بودید.

مهدی اشرفی گرامی، اگرچه دور و در ایران، ممنونم بابت مهربانیهات و پرکردن لحظهاتم با اشعار شگفت انگیزت. .بیصبرانه مشتاق خواندن کتاب جدیدت هستم.

دوست دوست داشتنی من **زهرا طاهری**، دوستی با تو یک سفر شگرف بی پایان است، بابت بودنت سپاسگزارم.

حمیرا ابراهیمی، **کاظم**، **آلما**، **امیرعلی** و **امیرمسعود جلالیپور**، مهر بی پایان شما تحمل دوری از وطن و خانواده را برای من قابل تحمل تر کرد. هیچگاه گرمای محبت شما را فراموش نخواهم کرد.

فرنوش خداداده، **پریسا نوروزی طلب**، **دیرک** و **کاریتک**، شما جزء بهترین انسان‌هایی هستید که من سعادت ملاقاتشون رو داشتم. برای من، شما تجسم هر چه خوبی هستید. سپاسگزارم بابت تمام مهربانی ها و دوستی تان.

بهاره سلیمی و **حمید حدادی**، شما در کنار خانواده ام عزیزترین و نزدیکترین انسانها به من هستین و بیشترین تاثیر رو در زندگی من گذاشتین، با شما خندیدم، گریه کردم، زندگی کردم و معنای عشق و دوست را تجربه کردم. حسِ خوشایندِ "داشتن یک دوست خوب "بیش از چیزیست که بشود توصیفش کرد. کلمات را یارای این نیست که بیان کنم چقدر وجود شما برایم عزیز و ارزشمند است.

و بالاخره می رسم به انسان‌هایی که بودنم از آنهاست، با یادشان و مهربانی هایشان آرام می گیرم و در هر لحظه از بودنم، سپاسگزار و عاشقشان هستم: **مامان**، **بابا**، **رادفه**، **فاطمه**، **مهدی**، **آرزو**، **اعظم**، **پریسا**، **بهدخت**، **نوروز**، **شهروز**، **سیدجواد**، **میلاد**، **محمد**، **آنا** (زهرا) و **امیر** گرمترین واژه های سپاسگزاری ام رو به شما پیشکش می کنم.

با مهر

سریه قربانی

فروردین یکهزار و سیصد و نود و سه